Effectiveness of Bonus and Penalty Incentive Contracts in Supply Chain Exchanges: Does National Culture Matter?

Yun Shin Lee Dina Ribbink Stephanie Eckerd

In this study, we investigate the impact of national culture on the effectiveness of bonus and penalty incentive contracts in supply chain exchanges. We conducted laboratory experiments in Canada, China, and South Korea, involving transactional exchanges in which suppliers were presented with either bonus or penalty contracts. Then we compared suppliers' contract acceptance, level of effort, and shirking across national cultures. Our findings reveal critical cultural influences on contract effectiveness. We show that although acceptance of bonus contracts is comparable across cultures, suppliers from Canada, a national culture considered low in power distance and high in humane orientation, exhibit lower acceptance rates of penalty contracts. In addition, we find evidence that suppliers associated with collectivist cultures exert more effort and shirk less in bonus contracts but these relationships also are more complex. When we compare contract effectiveness across bonus and penalty contracts within a given cultural setting, we find in all three countries greater acceptance of bonus contracts than penalty contracts. Also, after contracts are accepted, bonus contracts are more successful in China because suppliers exert greater effort and shirk less under bonus contracts than penalty contracts. However, in Canada and South Korea, the results of accepted contracts for both penalty and bonus contracts are nearly indistinguishable.

Keywords: behavioral operations; intra-cultural comparison; supply contract; incentives

1. Introduction

Contract effectiveness is a critical issue in supply chain exchanges. Contracts include details about the specifics of deliverables, set expectations, and provide performance evaluation tools (Ellram et al., 2004). Contracts are effective when they achieve efficient coordination of activities, fair outcomes, and safeguard against opportunism (Shou et al., 2016). Their role is especially critical in transactional exchanges, which are defined as arms-length exchanges between distinct firms and characterized by low interdependence and low asset specificity (Williamson, 1985, 1991). Transactional exchanges are typically associated with simpler, functional purchases (Sanders et al., 2007; Williamson, 1991) governed by contracts "composed of specific, short-term, and monetizable obligations entailing limited involvement of the parties" (Morrison and Robinson, 1997: 229). In the contexts of limited information sharing and restricted contract observability, different economic incentives can be embedded in these contracts as drivers for adherence and alignment, and therefore, improve overall contract effectiveness (Ketchen and Hult, 2007).

In practice, a variety of economic incentives can be applied to contracts to coordinate activities between supply chain entities, typically a buyer and a supplier. These incentive contracts fall into two broad types: bonuses and penalties. Bonus contracts generally have a fixed base amount at the inception of the contract as well as an additional reward to motivate the supplier's performance, discourage waste and inefficiency, and avoid moral hazard problems. For example, one type of bonus may be the renewal of the contract (Woods, 1999). In contrast, penalty contracts may specify monetary deductions from a base payout if it is determined that contract specifications are not met (e.g., late delivery, nonconformance with product specifications). In contracts for raw materials, the buyer often requires quality certification; if the supplier fails to meet the required specifications, then a financial penalty is imposed. Construction contracts regularly include both bonus and penalty clauses associated with project completion times (Lewis and Bajari, 2014).

Incentives may constitute a means for improving contract effectiveness; however, the degree of their success depends on *how* the recipients perceive the incentives. Previous empirical research has focused on

fairness, risk aversion, and opportunism as three key behavioral considerations known to influence the effectiveness of incentives. First, a wealth of previous research demonstrates that perceptions of organizational justice influence numerous work-related outcomes (Li and Cropanzano, 2009). Fairness, especially, has been leveraged to help explain the effectiveness of incentives in behavioral economics and social psychology. In a case of contracted effort between principals and agents, Fehr et al. (2007) show that in the aggregate inequity aversion motivates contract choice predictions, and Hannan et al. (2005) find evidence that employees increase their effort the fairer they perceive their contracts to be.

Second, risk aversion presents another behavioral consideration. The trade-off between risk and incentives has long been of interest in economics (Prendergrast, 2002) and management (Bloom and Milkovich, 1998). The literature in both fields is largely grounded in violations of the tenet of invariance, wherein framing effects lead to preferential decision-making in problems in which the choices are equivalent (Tversky and Kahneman, 1986). In particular, previous research shows that individuals exhibit risk avoidance when outcomes are framed as gains and are risk-seeking when outcomes are framed as losses (Kühberger et al., 1999). As for incentive contracts, gains are equivalent to bonus contracts and thus should be associated with risk aversion; penalty contracts represent losses and thus should be associated with risk-seeking behavior (Hossain and List, 2012).

Third, opportunism figures prominently in the literature on contracting. According to economic theory, one of the key benefits of contracting is to safeguard against a partner's opportunism (Eisenhardt, 1989; Williamson, 1985). However, previous theoretical work argues that formal contracts may actually encourage opportunistic behavior (Ghoshal and Moran, 1996). Empirical results on this front have been mixed and seem to suggest that the characteristics of the contract play into this relationship (Cavusgil et al., 2004). Contract framing (such as through bonus and penalty incentives) is one of those characteristics proposed to influence opportunistic behavior (Weber and Mayer, 2011).

In summary, the literature regarding fairness, risk aversion, and opportunism in contracting provides a strong rationale for the inclusion of these behavioral perceptions in the consideration of incentive contract effectiveness. However, there are also noted differences in how fairness, risk, and opportunism are

perceived across societies (Doney et al., 1998). This suggests an opportunity to delve into a more meaningful and nuanced perspective on how these behavioral considerations intersect with that of national culture in regards to incentive contracts. Indeed, prior research has noted the importance of congruence in organizational-societal culture toward achieving organizational legitimacy, a concept particularly relevant when organizational reward systems are concerned (Giacobbe-Miller et al., 2003). As such, the contribution of our research focuses on the interplay of the behavioral facets of fairness, risk, and opportunism with national culture to theoretically and empirically evaluate their impact on the effectiveness of incentive contracts.

In this research, we leverage the cultural dimensions of power distance, humane orientation, and individualism/collectivism — all culled from two well-established theoretical frameworks in Hofstede (1980) and GLOBE (House et al., 2004) — to develop novel hypotheses exploring the interplay of national culture and perceptions of fairness, risk, and opportunism on the effectiveness of incentive contracts. We empirically test our hypotheses using data collected from laboratory experiments conducted in three countries (Canada, China, and South Korea), selected for their cultural diversity. In these experiments, a buyer proposes either a bonus or penalty incentive contract to a culturally similar supplier. Effectiveness of the contract is then assessed in terms of suppliers' acceptance rates, the designated level of suppliers' effort exerted for these accepted contracts, and the rates at which suppliers shirk the terms of their contracts. We consider an incentive contract to be more effective if it induces a higher acceptance rate, a higher level of effort, and a lower shirking rate. With this experimental design, we can investigate a range of suppliers' decisions in response to buyers' contract offers within a variety of national cultures.

Our research makes several contributions to the literature on supply chain contracting. We theorize on the effectiveness of bonus and penalty incentive contracts in terms of important behavioral differences across national cultures. We argue that behavioral perceptions of fairness, risk, and opportunism are experienced differently across cultures and that these differences therefore have distinct impacts on how incentive contracts are received and responded to in transactional supply chain exchanges. Empirically, our findings show that the use of penalty contracts in cultures characterized by low power distance and

high humane orientation may lead to lower acceptance rates; on the other hand, in collectivist cultures suppliers may respond to bonus contracts with greater levels of effort and less shirking. Our investigation also considers the comparative effectiveness of bonus and penalty contracts within a given cultural context. This consideration allows us to suggest cultural settings in which certain contract types may be more successful. We show that in all three countries bonus contracts are accepted more than penalty contracts. However, after contracts are accepted, bonus contracts perform better in China because they elicit higher levels of effort while incurring fewer instances of shirking. In contrast, the results of bonus and penalty contracts are indistinguishable in Canada and South Korea in terms of both levels of effort and rates of shirking. Through our adoption of a cultural lens, we contribute to the burgeoning area of behavioral operations and supply chain research that attempts to address "how and why culture matters" in the contracting arena (Metters et al., 2010: 183). Our work has critical implications for both theory and practice.

2. National Culture Dimensions

Culture is defined as "the collective programming of the mind that distinguishes the members of one group or category of people from another" (Hofstede, 1991: 5). Previous studies in supply chain management have examined the impact of national culture on decision-making behavior in a variety of contexts. In a study of pharmaceutical manufacturing facilities, Gray and Massimino (2014) find that cultural congruence between a manufacturing plant and the organization's headquarters facilitates process compliance. Ribbink and Grimm (2014) show that cultural differences affected tactics in buyer-supplier negotiations and in an international context resulted in lower profits than were gained in a domestic context. Özer et al. (2014) identify cross-national differences toward trust and trustworthiness between the U.S. and China and show how these differences affect forecast sharing. The U.S. and China were also the focus of a study by Eckerd et al. (2016) in which they evaluate behavioral and attitudinal responses after a supply disruption caused by a psychological contract breach. Although these works demonstrate

the importance of national culture in the domain of supply chain exchanges, there remains an incomplete understanding of the cultural implications associated with the effectiveness of incentive contracts.

Researchers have developed several theoretical models to describe cultural characteristics and illuminate important differences between cultures. We draw on three cultural dimensions from the Hofstede (1980) and GLOBE (House et al., 2004) theoretical frameworks to inform our research: power distance, humane orientation, and individualism/collectivism (hereafter referred to as individualism). Power distance describes the extent to which a culture is accepting of hierarchy, authority, and status privileges. Individuals in cultures low on the power distance scale desire greater equality among members of society but those high on the power distance scale are more tolerant of inequities. Organizationally, power distance has implications for the legitimacy of employees' voices in workplace issues and decision making. This dynamic affects outcomes as varied as job performance, absenteeism, and satisfaction (Lam et al., 2002) as well as preferences towards conflict dispute resolution processes (Tyler et al., 2000). Of the countries included in our investigation, Canada ranks lowest on the power distance metric, while China and South Korea rank high. Both the Hofstede and GLOBE frameworks incorporate a measure for power distance.

Humane orientation describes societal emphasis towards fair action, altruism, generosity, kindness to others, and caring. Individuals from cultures higher on the humane orientation scale are likelier than those lower on the scale to encourage and even reward fair action. Humane orientation informs work across a variety of business contexts — for example, in project portfolio management (Unger et al., 2014) and corporate social responsibility performance (Miras-Rodriguez et al., 2015). This dimension is specific to the GLOBE framework, with no corresponding dimension in the Hofstede model. In accordance with the GLOBE study, Canada and China function similarly in humane orientation practices with values significantly above average, but South Korea falls below average.

Individualism, a dimension prominent in both theoretical frameworks, relates to the level of interdependence between members within a society and describes how individuals prefer to be rewarded for their efforts. High individualism implies self-reliance and reward systems based on individual merit;

low individualism (i.e., collectivism) indicates more interdependent and cooperative relationships focused on group goal achievement. Accordingly, individualism is defined as the degree to which individuals act in reference to themselves rather than a group. Differences along the dimension of individualism have been found to influence a range of organizational behaviors, such as the propensity to cooperate (Parks and Vu, 1994) and engage in organizational citizenship behaviors (Moorman and Balkely, 1995). As for the three countries studied in our research, Canada is described as a highly individualistic society, but China and South Korea are both considered collectivist cultures according to the Hofstede and GLOBE frameworks.

The values for power distance, humane orientation, and individualism for Canada, China, and South Korea are provided in Table 1. In the next section, we develop our experimental set-up employed in this research. It is also the launching point for our subsequent cultural argumentation and hypothesis development.

Table 1   National Cultural Dimension Values

	Power distance		Humane orientation	Individualism/collectivism	
	Hofstede	*GLOBE*	*GLOBE*	*Hofstede*	*GLOBE*
Canada	39	4.82	4.49	80	4.38
China	80	5.04	4.36	20	4.77
South Korea	60	5.61	3.81	18	5.20

Notes: There is no Hofstede correlate for humane orientation. Hofstede scores rank on individualism, with Canada scoring highest on individualism; GLOBE scores rank collectivism, with South Korea scoring highest on collectivism. The Hofstede dimensions are assessed on a 0-100 (low-high) point scale; GLOBE dimensions are assessed on a scale of 1-7 (low-high).

3.   Incentive Contract Model and Hypotheses Development

Our research setting involves a buyer and supplier with conflicting goals involved in a transactional exchange in which each party desires to maximize their profit. It is also common in these supply chain exchanges troubled by misaligned goals that buyers use reward and penalty incentives to help them align their interests with the suppliers (Ketchen and Hult, 2007). When such economic incentives are used, a buyer, because of problems of moral hazard and adverse selection (Eisenhardt, 1989), has difficulty

verifying the actions of the supplier; hence, costly verification processes are necessary to ascertain the supplier's true degree of effort. As agency theory captures these aspects of our research setting, we use a simple principal-agent problem, adapted from Fehr et al. (2007), to define the bonus and penalty incentive contract models within the context of a transactional supply chain exchange. We then use these contract models to set up our hypotheses and experiment.

Our focus is on gaining a better understanding of the underlying cultural influences driving the decisions made by suppliers as agents. A buyer proposes a contract to a supplier to perform a production task. In all contracts, the buyer specifies a fixed fee w and a demanded effort level e^*. Let $\underline{e}$ be the minimum effort level the supplier could choose. If the supplier expends effort $e \geq \underline{e}$, they generate a gross profit $v(e)$ for the buyer that is strictly increasing and concave in e, but they also incur a private cost $c(e)$ that is strictly increasing and concave in e. We assume that $c(\underline{e}) = 0$.

The buyer also states an incentive, either a bonus or a penalty, in the contract. The buyer pays a bonus or imposes a penalty only when a third party, i.e. the courts, verifies the supplier's effort level and enforces the contract exogenously (Fehr and Gächter, 2000; Roels et al., 2010). The parties to a contract cannot perfectly enforce its terms. This makes third-party verification and enforcement necessary; even if both parties observe an actual effort level, without enforcement a buyer may have reason not to pay a bonus and a supplier may have reason not to pay a penalty. For example, a buyer or supplier may be in a difficult financial situation and unable to make a bonus or penalty payment, respectively. Indeed, in practice, there are many legal disputes about unsettled incentive payments. For example, Sun et al. (2013) discuss the legal issues of incentive contracts in transportation. To summarize, we assume that both parties observe a supplier's effort, but to contract on effort using incentives, the courts should be able to verify the level of effort and enforce the conditions of the contract (Fehr et al., 2007). The courts enforce only those contractually specified payments that are contingent on verifiable events. If the level of effort cannot be verified, then the contracted incentive payment is not enforceable by a court of law.

In many cases, the courts can verify supplier's effort level only if the buyer invests in verification technology (e.g., a system that monitors and reports effort). We assume that this verification technology renders effort verifiable to the court with probability p (Fehr and Schmidt, 2007) and that it requires a fixed cost, partly because of information stickiness (Xue and Field, 2008). Thus, the probability of verification can be interpreted as the probability of legal enforcement of the incentive contract (Kvaløy and Olsen, 2016) and this leads to *probabilistic* bonus and penalty payments.

Reliance on the verification process (i.e., probability p) depends on the nature of the requested effort. If the effort pertains to increasing the quality of production of some good or service, it is neither easily measurable nor verifiable. Consequently, the reliability of the verification process remains uncertain (i.e., p is small). For example, if a supplier invests effort in maintaining machinery, it would be prohibitively costly to even attempt to verify the effort level. Also, in raw material contracts, buyers regularly include a requirement for chemical composition tests to ensure that the mix of the product matches the contract specifications. A subsample of the procured product, not the complete shipment, is usually analyzed, which makes the reliability of verification process uncertain. If the effort relates to the quantity of production, then it is much easier to measure and verify the effort, and the verification process will be more reliable (i.e., p is high) and legal enforcement of the incentive contract will be more certain (Fehr and Schmidt, 2004). When $p=1$, the verification process is certain, but it is still necessary to enforce contractual terms.

In our experimental setup, we assume that the verification process is uncertain, and the buyer is required to invest a fixed cost k that allows for partial verification of the supplier's effort level with probability p, $0 < p < 1$. In other words, we let the probability of legal enforcement of the incentive contract be equal to p. We also set this same probability across the bonus and penalty contracts; in both contracts, the null hypothesis is set at correctly identifying incentive payment events. The null is accepted with p chance, and the null is incorrectly rejected with $(1-p)$ chance. This means that the likelihood of correctly identifying a case of paying a bonus in the bonus contract is the same as that of incurring a penalty in the penalty contract. Note that probability p is in the verifiability of the effort because of the

degree of reliance on the verification process and not on the ability to observe the actual level of effort. Under the *Bonus Contract,* the buyer awards a bonus b to the supplier if non-shirking $(e \geq e^*)$ is verified with probability p. However, the supplier needs to be offered a minimum bonus, i.e., the bonus b is bounded below by $\underline{b}$. Under a *Penalty Contract,* the buyer imposes a penalty (fine) f on the supplier if shirking $(e < e^*)$ is verified with probability p. However, the supplier cannot be punished arbitrarily harshly, i.e., the penalty f is bounded above by $\bar{f}$. If the supplier rejects the contract, both buyer and supplier get their reservation utilities that are equal to λ (e.g., a payoff from an outside option).

3.1. The Timing of Events and Mechanisms

We used the subscripts B and P, respectively, to designate the bonus and penalty contracts. Similarly, we use the superscripts B and S, respectively, to indicate buyers and suppliers.

3.1.1. Bonus Contract.
At date 0, the buyer incurs a fixed cost k to contract on effort and offers a contract (w_B, e_B^*, b) that stipulates a fixed fee w_B, a demanded effort level e_B^*, and a bonus b, that the buyer pays to the supplier if non-shirking is verified. At date 1, the supplier first decides whether to accept or decline the contract and then chooses an effort level e_B upon acceptance. At date 2, with the probability p, the supplier's effort level is verified. Then payoffs are realized and payments made. The buyer's expected payoff is defined by

$$M_B^B = v(e_B) - w_B - k - pb \text{ if } e_B \geq e_B^*$$

$$M_B^B = v(e_B) - w_B - k \quad \text{if } e_B < e_B^*$$

The supplier's expected payoff is given by

$$M_B^S = w_B - c(e_B) + pb \quad \text{if } e_B \geq e_B^*$$

$$M_B^S = w_B - c(e_B) \quad \text{if } e_B < e_B^*$$

3.1.2. Penalty Contract.
At date 0, the buyer incurs a fixed cost k to contract on effort and offers a contract (w_P, e_P^*, f) that stipulates a fixed fee w_P, a demanded level of effort e_P^*, and a penalty f, to be paid by the supplier to the buyer if shirking is verified. At date 1, the supplier first decides whether to accept or decline the contract and then chooses an effort level e_B upon acceptance. At date 2, with the

probability p, the supplier's effort level is verified. Then payoffs are realized and payments made. The buyer's expected payoff is defined by

$$M_P^B = v(e_P) - w_P - k \quad \text{if } e_P \geq e_P^*$$

$$M_P^B = v(e_P) - w_P - \text{k} + pf \text{ if } e_P < e_P^*$$

The supplier's expected payoff is given by

$$M_P^S = w_P - c(e_P) \quad \text{if } e_P \geq e_P^*$$

$$M_P^S = w_P - c(e_P) - pf \quad \text{if } e_P < e_P^*$$

In Appendix A.1, we provide an analysis of the optimal contract design in which each player is self-interested in maximizing their expected payoffs.

3.2. Hypotheses

We investigate the influence of culturally embedded behavioral differences on the effectiveness of incentive contracts (either bonus or penalty) based on suppliers' decisions. We first theorize regarding the impact of culture on a supplier's decision to accept a contract from a buyer, followed by the level of effort the supplier delivers. We conclude with an evaluation of shirking by suppliers across cultures.

3.2.1.Supplier Acceptance of Contract. The first decision a supplier makes when a buyer presents a contract is whether to accept or reject it. This decision can be linked to considerations of organizational justice, or to how fair the decision maker perceives the incentive contract to be. Broadly speaking, research in supply chain management has demonstrated how justice considerations can greatly improve buyer-supplier relationships (Liu et al., 2012; Narasimhan et al., 2013). When incentive contracts are involved, bonus contracts generally are perceived as fairer than penalty contracts. The rationale underlying this is that bonuses are extensions of kind or friendly behavior that in turn produces reciprocal behavior (Fehr and Falk, 2002; Fehr and Gächter, 2000).[1] As such, acceptance rates tend to be greater under bonus framing (Fehr and Gächter, 2000; Wuttke et al., 2013).

---

[1] Bonus contracts may also be perceived as fairer than penalty contracts when individuals use a mental accounting rule called prospective accounting in which coupling is stronger when looking forward in time, but weaker when looking backward in time (Chen et al., 2013). Under incentive contracts, individuals may mentally segregate fixed

However, previous research has not considered cultural differentiation in acceptance rates. This is a notable gap, because previous research has shown significant between-group differences regarding notions of fairness (Buchan and Croson, 2004) and punishment resulting from distinctions in social institutions (Henrich et al., 2005). As such, it is important to understand the influence of fairness on perceptions of contract framing across cultures. The cultural dimension of power distance serves as a useful theoretical lens for evaluating the interplay of culture and fairness on a supplier's decision to accept or reject contracts. Power distance is relevant in that buyers impose contracts, and a supplier has no voice in the development of contract terms. The notion of voice is a crucial determinant in perceptions of procedural justice; lower levels of voice tend to provoke unfavorable reactions from people in cultures with low power distance (Brockner et al., 2001). As such, we argue that because buyers control the process, suppliers' acceptance rates will be noticeably lower for both bonus and penalty contracts in cultures characterized by low power distance, the opposite of cultures characterized by high power distance.

However, we do not expect differences in power distance to operate in isolation. Cultural variations in fairness perception are also incorporated within the humane orientation dimension of the GLOBE study (House et al., 2004). As such, this dimension may be useful in theoretically discerning additional cultural nuances in a supplier's decision to accept or reject contracts. Humane orientation describes societal emphasis on fairness; high humane orientation cultures encourage and even reward fair actions. We therefore argue that individuals associated with high humane orientation cultures should put greater emphasis on the fairness of a bonus contract and reward their issuers accordingly with higher acceptance rates. Similarly, people from high humane orientation cultures should disapprove of unfair penalty contracts and punish their issuers with lower acceptance rates. When taken in conjunction with the power

---

payments and incentive payments, respectively, because these payments are made before and after the verification process. Then, under the bonus contract, having a lower fixed payment is strongly coupled with the possibility of receiving a bonus, whereas the verified bonus payment is weakly coupled with the earlier lower fixed payment. Under the penalty contract, having the higher fixed payment is strongly coupled with the possibility of incurring a penalty, whereas the verified penalty payment is weakly coupled with the earlier higher fixed payment. With prospective accounting, bonus contracts can therefore be perceived to be fairer.

distance dimension, we propose that a culture with low power distance combined with high humane orientation (i.e., Canada) will exhibit significantly lower acceptance rates under the framing of penalty contracts. This result is because both cultural effects reinforce one another directionally – both should have the effect of decreased contract acceptance. However, those individuals associated with a culture high in power distance and low in humane orientation (i.e., South Korea) should be generally less attuned to issues of justice and fairness, thereby exhibiting a comparatively higher acceptance rate of penalty contracts. Alternatively, we expect that under bonus contracts, the cultural effects of power distance and humane orientation will offset each other, leaving no discernible difference in the acceptance rates across the cultures included in our study. Table 2 summarizes the directionality of these arguments. In summary, fairness considerations are crucial to acceptance decisions, and because fairness is reflected in both the power distance and humane orientation dimensions of culture, we state the following hypotheses:

H1A: *There will be no significant difference in the acceptance rates of bonus contracts across cultures.*

H1B: *Individuals from cultures associated with low power distance and high humane orientation will exhibit significantly lower acceptance rates under penalty contracts than those from high power distance and low humane orientation cultures.*

Table 2   Comparison of Supplier Acceptance Rates across Cultures

	Cultures associated with low power distance and high humane orientation (Canada)			Cultures associated with high power distance and low humane orientation (Korea)		
	Low power distance	High humane orientation	Overall	High power distance	Low humane orientation	Overall
Effect on acceptance rate under bonus contract	-	+	Balance out	+	-	Balance out
Effect on acceptance rate under penalty contract	-	-	Low acceptance rate	+	+	High acceptance rate

3.2.2. Supplier Level of Effort. We next seek to understand the influence of culture on the level of effort a supplier expends on a contract. Once a contract is accepted, a supplier must determine how much

effort to expend, recognizing that the decision involves risk. Due to the uncertainty of the verification process, suppliers who exert effort equal to or higher than what is demanded under the bonus contract receive a *probabilistic* bonus (i.e., there is a risk of not getting a bonus). Similarly, under the penalty contract, suppliers who exert less effort than what is demanded pay a *probabilistic* penalty (i.e., there is a risk of paying a penalty). Previous research has revealed that the degree of risk acceptable to an individual is theoretically tied to the cultural dimension of individualism (Breuer et al., 2014; Hsee and Weber, 1999; Weber and Hsee, 1998). Cultural stereotypes suggest that people in Western cultures are more risk-seeking than East Asians, which is largely attributable to Westerners' heightened sense of self-esteem and control (Breuer et al., 2014; Heine et al., 1999; Peng and Nisbett, 1999). However, a number of studies present empirical evidence and arguments for an opposing view. Meisel et al. (2016) observe that the Chinese participants in their study are more risk-seeking than their U.S. counterparts. Hsee and Weber (1999) evaluate individuals' choices between risky options and sure outcomes in the U.S. and China and find evidence that East Asians are more risk-seeking than Westerners in financial matters. They explained this observed behavior in terms of individualism. As a collectivist culture, East Asians demonstrate less risk aversion in financial situations because of a "cushion" afforded them via their in-group social networks. This social cushion makes East Asians less afraid of financial risks because they expect family or other in-group members will step in to help any group member who selects a risky option and subsequently encounters a large and possibly catastrophic loss. In contrast, individualistic cultures expect individuals to bear the consequences of their own risky financial decisions, thereby tempering their likelihood of accepting risky options (Hsee and Weber, 1999).

We extend this logic to suppliers' decisions regarding how much effort to exert under bonus and penalty contracts. When an individual increases his effort under a bonus contract, there is the salient risk of not getting the bonus because of the uncertain verification process. Therefore, we expect that under the framing of a bonus contract, less risk averse individuals associated with collectivist cultures (i.e., China and South Korea) will exert greater effort. Alternatively, under a penalty contract, an individual who decreases their level of effort risks paying a penalty. Therefore, we expect that within the framing of a

penalty contract, less risk averse individuals from collectivist cultures will under these circumstances exert less effort. In summary, decisions surrounding a supplier's level of effort are inherently risky, and because risk is manifest within the individualist dimension of culture, we hypothesize the following:

H2A: *Individuals from collectivist cultures will exert significantly more effort under bonus contracts than those from individualist cultures.*

H2B: *Individuals from collectivist cultures will exert significantly less effort under penalty contracts than those from individualist cultures.*

3.2.3. Supplier Shirking. Finally, the notion of costly effort combined with probabilistic realization of penalties or rewards provides some impetus for a supplier to behave opportunistically. Consistent with the principal-agent literature involving moral hazards (Alchian and Demsetz, 1972; Eisenhardt, 1989), we define the supplier's opportunistic behavior as *shirking* (Aron et al., 2005; Handley and Benton, 2012). Although shirking is clearly related to a supplier's level of effort, it indicates deliberate underperformance, meaning the supplier provided less effort than the buyer's contract demanded. This distinction is important, as such we evaluate shirking separate from level of effort. The previous literature has shown that the average effort suppliers expend is much less than the average effort buyers demand under penalty contracts (Fehr et al., 2007). However, here again we expect that cultural individualism influences whether a supplier decides to shirk or deliver the effort level a buyer requires.

Previous literature has tied the dimension of individualism to patterns of cooperation in situations described as a social dilemma, or those situations in which the parties have competing interests (Parks and Vu, 1994), specifically finding that those from collectivist cultures cooperate more. Whereas Parks and Vu (1994) examine outcomes under a public goods game, the buyer-supplier contractual domain also presents a conflict situation in that each party wants to maximize profits, often at the expense of the other. Indeed, shirking represents an extreme manifestation of profit-seeking by a supplier. However, opportunistic profit-seeking in this manner is highly inconsistent with a collectivist culture in which self-interested behaviors are avoided, unlike in individualist cultures that value individual gain. Further, Luo (2007) find that in a collectivist context, social forces serve to stifle opportunistic behavior, but an

individualist context requires more formal forms of governance for the same purpose. We extend this logic to theoretically argue that shirking behavior is less likely by individuals from collectivist cultures than those from individualist cultures. We expect this pattern to hold across both bonus and penalty contract frames. In summary, cooperativeness is an important factor in a supplier's decision to shirk, and because cooperativeness is reflected within the individualism dimension of culture, we hypothesize the following:

H3: *Individuals from collectivist cultures will shirk significantly less than those from individualist cultures.*

4. Experimental Design and Procedure

4.1. Experimental Design

The buyer requests an effort level (demanded effort) $e \in \{1, \dots, 10\}$ from the supplier. The supplier can choose an effort level $e \in \{1, \dots, 10\}$ associated with a corresponding effort cost (see Table 3). Each effort level corresponds to a gross profit $v(e) = 10 \cdot e$ realized by the buyer. To contract on effort, the buyer invests a fixed cost $k = 10$ that allows for verification of the supplier's effort: an assessment of how the level of effort the buyer demands corresponds with the effort the supplier delivers. A supplier "shirks" when his level of effort is less than what the buyer demands. The supplier can receive a bonus when the effort is verified as equal to or greater than the level of effort demanded in the bonus treatment. The supplier can incur a penalty when the effort is verified as less than the level of effort demanded in the penalty treatment. We assume that the verification process is uncertain with probability p=1/3. The maximum penalty/minimum bonus the supplier can be charged/rewarded is predetermined at 12. These boundaries are chosen to make the two incentive contracts economically equivalent (the expected payoffs of the supplier and buyer in equilibrium are the same in the two contracts; please refer to Appendix A.2 for more details). The buyer is constrained to choose a fixed fee $w \geq c(e^*)$ in the penalty contract and $w \geq c(e^*) - pb$ in the bonus contract, which avoids losses for the supplier if, on average, they meet the

buyer's demands for level of effort. If the supplier rejects the offered contract, both parties get their reservation utilities $\lambda = 1$ (e.g., a payoff from an outside option). The flow chart in Figure 1 shows a description of the steps and decisions in the experiment. Given these parameters, we can obtain the optimal contract for the buyer, optimal effort level for the supplier, and payoffs in equilibrium when both players are self-interested as provided in Table 4 (refer to Appendix A.1 for the analysis of optimal contract design).

Note that the highest incentive-compatible effort level of the supplier is $c(e_B^*) = p\bar{b}$ in the bonus contract and $c(e_P^*) = p\bar{f}$ in the penalty contract (again, refer to Appendix A.1 for more details). In a situation of a completely uncertain verification process ($p = 0$), the task is non-verifiable in any case and a rational buyer would choose a minimum effort (such that $c(e) = 0$). In a situation with a completely certain verification process ($p = 1$), the task is perfectly verifiable in all cases. In this instance, a rational buyer would choose an effort equal to the bonus or penalty so that $c(e) = b \: or \: f$ (this can be obtained by substituting $p = 1$ in the highest incentive-compatible effort level equations stated above, i.e., $c(e_B^*) = p\bar{b}$ in the bonus contract and $c(e_P^*) = p\bar{f}$ in the penalty contract). When verification is uncertain with $0 < p < 1$, the optimal effort level for a buyer increases with the value of p, again as can be seen from the highest incentive-compatible effort-level equations. An experimental design that adopts completely certain or uncertain verification processes is less applicable to our study because we examine how the supplier's risk-taking behavior in the face of unknown outcomes affects his effort decisions (as we hypothesize in H2). Hence, we look at a case with some uncertainty in the verification (i.e., $0 < p < 1$). For a small p, the verification process is more uncertain, and the supplier's decision involves more risk. In this case, the difference in the supplier's level of effort because of cultural differences in attitudes toward risk will be more pronounced. As the value of p increases, this difference becomes negligible. Therefore, we assume a small p ($=1/3$) in the experiment. Importantly, this value is consistent with that used in other incentive contract experiments in the behavioral economics literature (e.g., Fehr et al., 1997; Fehr et al., 2007; Fehr and Schmidt, 2007). Moreover, a change in the value of p is unlikely to

significantly change participants' responses. Bigoni et al. (2012) note that a person's perception of an event occurring is not based on actual probabilities but reflects other factors such as the size of the penalty. In addition, cognitive biases often cause participants to disregard the actual probabilities and focus on other more tangible aspects of a contract (Simon et al., 2000; Tversky and Kahneman, 1974).

Our experimental design adheres to an abstract representation of the buyer-supplier relationship so that we may generate insights applicable to most practical relationships. We intentionally avoided using a more detailed supply chain context because we did not want to introduce context-specific bias. In this regard, our paper shares a similar spirit with other papers that adopted an abstract setting applicable to a wide spectrum of supply chain contexts (for example, see the two-person sequential mover game in Loch and Wu (2008) and the two-person social dilemma game in Wu et al. (2011)).
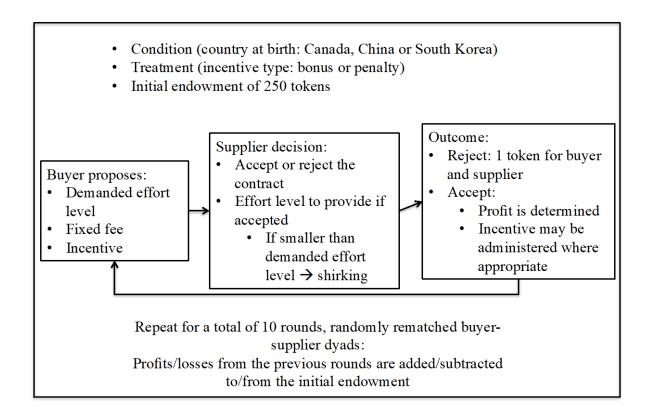
Table 3   Supplier's Effort Cost Function

e	1	2	3	4	5	6	7	8	9	10
c(e)	0	1	2	4	6	8	10	13	16	20

Table 4   Optimal Contract and Payoffs in Equilibrium When Both Players Are Self-Interested

	Bonus contract	**Penalty contract**
Optimal contract for the buyer	$(w_B, e_B^*, b) = (1,4,12)$	$(w_P, e_P^*, f) = (5,4,12)$
Optimal effort level for the supplier	$e_B = 4$	$e_P = 4$
Payoff in equilibrium	$M_B^S = 1$ $M_B^B = 25$	$M_P^S = 1$ $M_P^B = 25$

Figure 1 Flow Chart of Experiment

- Condition (country at birth: Canada, China or South Korea)
- Treatment (incentive type: bonus or penalty)
- Initial endowment of 250 tokens

Buyer proposes:
- Demanded effort level
- Fixed fee
- Incentive

Supplier decision:
- Accept or reject the contract
- Effort level to provide if accepted
  - If smaller than demanded effort level → shirking

Outcome:
- Reject: 1 token for buyer and supplier
- Accept:
  - Profit is determined
  - Incentive may be administered where appropriate

Repeat for a total of 10 rounds, randomly rematched buyer-supplier dyads:
Profits/losses from the previous rounds are added/subtracted to/from the initial endowment

4.2. Experimental Procedure

We used a 3 (country: Canada, China, and South Korea) x 2 (incentive structure: bonus and penalty) between-subject design resulting in six factorial variations. We will refer to variations in country as experimental conditions and to variations in the incentive structure as experimental treatments. The experiments were conducted in computer laboratories in each country in the respective language of each location. All the original materials for the experiments were developed in English and then translated and back-translated into the other languages by native speakers to assure consistency. Across all experiments, a native speaker recorded the instructions that were then played back during the experiments, thus achieving consistency across treatments and sessions. All sessions were conducted with one author and a local aide at each location, with either or both of those individuals being a native speaker and experienced experiment administrator.

Each session consisted of three parts: an experiment (choice game) based on Zhang et al. (2016), a survey to gather supplemental information, and the buyer-supplier contracting experiment (contract game). The first experiment, executed in zTree (Fischbacher, 2007), was used to determine participants' relative level of loss aversion. Next, participants were asked to complete a questionnaire that provided us with demographic information about the participants (including gender, educational background, and age) as well as their risk attitudes in specific areas such as investments and gambling (scales based on Weber et al. (2002)). Data obtained from the first experiment and the survey are used to develop control variables in our subsequent regression analysis.

The final part of each session was the buyer-supplier contracting experiment developed in the zTree software. During any experimental session, only one treatment was administered. Participants listened to the instructions and also received a hard copy of the instructions printed in their native language. To take part in the experiment, participants had to pass a 10-question quiz used to confirm their understanding of the experiment. Only after they answered all ten questions correctly were they allowed to continue with the actual experiment. Participants were only informed about the contract type assigned to them for that session; a participant in the bonus contract treatment was not given any information about penalty type contracts, and vice versa.

Each treatment involved between 36 and 42 participants. Sessions consisted of six, eight, or ten participants, depending on the number of participants that showed up for a given session. Within these groups, participants were randomly assigned the role of either buyer or supplier. Participants played their assigned roles for ten rounds and were randomly and anonymously matched with a new partner for each round. We made sure participants were aware of this. Thus, although participants might have known each other before entering the session, they did not know who they were matched with during any given round of the experiment. In addition, random re-matching after each round was necessary because our contract model assumes a one-shot game, and it also avoided reputation effects caused by familiarity with any one partner. Participants were not allowed to talk with one another during the experiment. All interactions with their partners were through standardized decisions on their computer screens.

All payouts in the experiment were calculated in tokens. The participants were told in the instructions the exchange rates of tokens into their local currency. All participants received an initial endowment (in tokens) before starting the experiment. Their individual endowments increased or decreased based on their own and their partners' decisions during the experiment. Instructions included a clause that would expel from the experiment any participant who lost a specified fraction of the initial endowment. This was to ensure that in all treatments participants could only earn positive amounts for their involvement; however, no one was expelled from any session. At the end of each of the ten rounds, participants were informed of their own and their partner's payouts and asked to assess the level of fairness of that round's contract. At the end of the experiment, participants were informed of their total payout and paid privately in cash before leaving the lab.

A critical aspect of the experiment involved achievement of compensation parity across countries. After the input of local experts, we adopted the Big Mac standard (Ong, 1997) to develop currency conversion rates for the experimental tokens. The initial endowment (exchange rate) in Canada was Can$12 (1 token = Can$0.048), in China ¥36 (1 token = ¥0.14), and in South Korea ₩8,200 (1 token = ₩32.8). The average payout (including initial endowment) in Canada was Can$14.01 (with a minimum of Can$12 and a maximum of Can$22), ¥43.13 in China (with a minimum of ¥36 and a maximum of ¥68), and in South Korea ₩9,986.88 (with a minimum of ₩8,200 and a maximum of ₩13,915). All payouts were rounded up to the nearest whole unit of currency. The minimum payout was bounded at the level of the initial endowment, although participants were not made aware of this until the experiment concluded. The entire experiment averaged about one hour in duration for each session.

4.3. Participants

The experiments were conducted at a large top 10 university in each country. To assess eligibility, participants were first screened by country of birth and educational background. All participants were natives of their respective countries (naturalized citizens were excluded). Participants were also required to have an educational background in science, engineering, or business. Across the three countries, 226 students in 113 buyer-supplier dyads completed the experiment (Canada: 74 participants – 37 dyads,

China: 72 participants – 36 dyads, South Korea: 80 participants – 40 dyads). In total, 114 participants (57 dyads) were assigned to the bonus treatment and 112 (56 dyads) to the penalty treatment. The average age of participants (assuming the roles of both buyers and suppliers) was 20.60, and 41% were female.

Our experiment used students. Therefore, it is important to justify this design decision. Stevens (2011) posits that the use of student samples should be dependent on the research question and parameters tested. Bolton et al. (2012) demonstrate in a newsvendor experiment that students perform much like experienced managers. They conclude that in a profit-maximizing context, students and managers appear to exhibit the same non-optimal behavioral tendencies, indicating that these biases are inherent in the participants, irrespective of their level of experience. Given the profit-maximizing context of our scenario, the same logic can be applied to our study. Similarly, students have been used in other research involving contract games — for example, Katok and Wu (2009) within a supply chain contract setting and Fehr et al. (2007) in a general principal-agent contract setting. Also, Özer et al. (2014) use a student sample to compare the trust, trustworthiness, and strategic information-sharing behavior of supply chain members' countries of origin (China versus the United States) in a cross-country supply chain experiment.

We also assume that participants native to and residing within a culture are indeed representative of that culture. This is consistent with a basic tenet in psychological research: "[I]ndividuals generally act in ways that correspond to cultural influences and expectations" (Berry et al. (1992), as cited in Berry (1997: 6)). Research has also demonstrated a timeframe of about six years as necessary for acculturation to a different culture, and these findings apply even to immigrant youths (Berry et al., 2006). The participants in our study, averaging over 20 years in age, were all born and raised in the respective countries in which we conducted the experiment. Therefore, we are confident that our sample is representative of the national cultures under study.

4.4. Measures

Our dependent variables are *acceptance rate, effort level,* and *shirking*, all based on suppliers' decisions. *Acceptance rate* is a binary variable and measures whether the supplier accepts the contract offered by the buyer (accept=1). *Effort level* is the actual level of effort the supplier provides to the buyer and is

measured along an ordinal scale from 1 to 10. *Shirking* is again a binary variable and measures whether a supplier's level of effort is less than what the buyer demanded (shirk=1). Note that although shirking is negatively correlated with effort level, the two variables - *effort level* and *shirking* - measure different things; *effort level* measures the effort level provided by the supplier, and *shirking* measures effort level conditional on demanded effort.

The main treatment variables are *country* and *treatment*. We created dummy variables for each of the three countries. Canada is used as the reference category in our analysis. The coefficient value of the country variable is the key to understanding the hypothesized cultural effects. *Bonus treatment* is a binary variable indicating whether the contract offered to the supplier includes a bonus or penalty incentive (bonus=1). The supplier's decisions (acceptance rate, effort level, and shirking) are also closely related to the three contract parameters that the buyer chooses (i.e., *fixed fee*, *demanded effort level,* and *incentive*), and so these are also included as control variables. The buyer pays an unconditional *fixed fee* to the supplier; at a minimum, this fee needs to cover the cost of the effort level demanded. *Demanded effort level* is the effort level requested by the buyer (between 1 and 10). Each effort level is associated with a different cost to the supplier (see Table 3). *Incentive* is the amount the buyer offers that reflects the bonus or penalty, depending on the assigned treatment. The model also includes a *time trend* to account for learning over the course of the experiment. Finally, we control for demographic factors so that the observed main effects are indeed because of cultural differences and not the demographics of the participants. We control for age (in years) and gender (male=1) in the regression analyses. In Section 5.4, as robustness checks, we introduce additional controls for individual differences in loss aversion, in equity aversion, and risk attitude.

Table 5 contains descriptive statistics of our key variables and the correlations between them for all 113 participants who assumed the role of the supplier during the experiment. We include only the supplier participants' data because their data include both the contract parameter values offered by the buyers' as well as the suppliers' decisions. Among the correlations for the contract parameters (fixed fee, demanded effort, incentive), participant controls (loss aversion score, risk attitudes, age, gender), and dependent

variables (acceptance, effort, and shirking), the highest correlation of 0.54 is between demanded effort and fixed fee, which is not unexpected and related to the nature of the experiment. As the demanded effort level increases, the offered fixed fee should also increase. There are also high negative correlations between shirking and effort level and between acceptance and effort level. However, these are all dependent variables not included in the same model. Table 6 contains separate comparisons of actual contract parameters and effort decisions versus optimal decisions (as shown in Table 4) for each treatment and each country. These comparisons show that actual decisions deviate significantly from optimal decisions.

5.   Experimental Results

Tables 7 and 8 present the results of the main analysis. In Table 7, regression results for the key dependent variables (supplier acceptance, effort, and shirking) are presented separately *by treatment* (bonus and penalty) for ease of discussion. Instead of reporting the actual regression coefficients, we report *odds ratios* to make the interpretation of the results more meaningful. Models I through VI show the results of the bonus treatment and models VII through XII show the results for the penalty treatment. Table 7 is used to test our hypotheses. In Table 8, we present the marginal means based on the main regressions reported in Table 7. Table 7 also presents the results from the robustness checks discussed in Section 5.4. Table 9 presents separate regression results for the same dependent variables *by country condition* (Canada, China, and South Korea) while controlling for the treatment (*bonus* or *penalty)*. The marginal means for the models presented in Table 9 are reported in Table 10. Tables 9 and 10 present a more nuanced view of the effects of each contract type within each country.

5.1.   Contract Acceptance

We first examine the results regarding contract acceptance. Using logistic regression analysis, we test the effect of a country's culture on the acceptance rates for bonus and penalty contracts while controlling for the contract parameters of fixed fee, effort level demanded, and incentive (bonus or penalty), supplier's gender and age, and a time trend. We hypothesized in H1a that under bonus contracts no discernible

difference would occur in acceptance rates across cultures. Our results, shown in Model I of Table 7, demonstrate no statistically significant differences in the acceptance rates between the three cultures, thus supporting H1a. This finding is further supported by a comparison of marginal means in Table 8 that shows no statistical difference in the level of acceptance rates between Canada, China, and South Korea. As for penalty contracts, we hypothesized (H1b) that individuals from cultures associated with lower power distance and higher humane orientation would exhibit lower acceptance rates. From Table 7, Model VII, we observe that Canadian participants, who possess the lowest power distance and the highest humane orientation, accepted significantly fewer contracts than South Korean participants ($p<0.05$), who are associated with a comparatively high score on power distance and the lowest score on humane orientation. With all factors kept constant, South Koreans were about 2.4 times as likely to accept a contract as their Canadian counterparts. However, no statistically significant difference is observed between Canadian and Chinese participants. We note that China, with high scores for both power distance and humane orientation, is an interesting case. Indeed, we see the impact of this reflected in the marginal means of Table 8 in which Chinese participants' acceptance rates under penalty contracts fall between those of Canadians and South Koreans. Altogether, these results lend support to H1b.

Several of our control variables exerted some influence over the acceptance decisions as well. We find in our results by treatment in Table 7 that for both the bonus and penalty contracts, the odds ratio of a fixed fee is significantly greater than 1 ($p<0.01$). Hence, participants are likelier to accept a contract if they are assured a greater fixed payoff. Based on the individual country regressions (Table 9), we show this is consistent across all three countries. Results by treatment also show that the level of effort demanded has a statistically significant negative effect on acceptance rates under both the bonus and penalty treatments ($p<0.05$ and $p<0.01$, respectively); higher requested levels of effort (and therefore higher potential costs to the supplier) are associated with lower acceptance rates. Again, this finding is consistent across countries. Referencing the results by treatment, we find the amount of the penalty is statistically significant and negatively related to acceptance rates in the penalty treatment ($p<0.01$), but no corresponding effect is found in the bonus treatment. Table 9 shows participants in Canada and China, but

not South Korea, drive the effect of the penalty parameter. Across both incentive treatments we find a positive time trend in our data ($p<0.05$ for bonus and $p<0.01$ for penalty), indicating that acceptance rates generally increased as participants gained experience with the game. Finally, we observe that as age increases, the likelihood of acceptance decreases only in the penalty treatment ($p<0.01$). Overall, our regression by treatment (Table 7) correctly classified 87% of the observations in the bonus and about 85% in the penalty treatments, indicating a reasonably good fit of the models. Similar levels of fit can be observed for the regression by country (Table 9).

Although within-country differences between the bonus and penalty contracts were not hypothesized, Table 10 shows that in all three countries bonus contracts are accepted at significantly higher rates than penalty contracts.

Table 5: Descriptive Statistics and Correlation Table for Suppliers

	Mean	SD	I	II	III	IV	V	VI	VII	VIII	IX	X
Shirking (I)	0.554	0.497	1									
Effort (II)	2.960	2.609	-0.782	1								
Acceptance (III)	0.835	0.372	-0.400	0.505	1							
Demanded effort (IV)	5.523	1.977	0.179	0.240	0.117	1						
Fixed fee (V)	16.053	11.100	-0.120	0.372	0.299	0.541	1					
Incentive (VI)	13.411	6.214	-0.053	0.125	0.018	0.260	-0.096	1				
Loss aversion score (VII)	1.682	1.412	-0.012	-0.033	0.022	-0.051	-0.012	0.010	1			
Investment risk (VIII)	-0.003	0.965	0.004	-0.012	0.005	-0.013	-0.001	0.055	-0.063	1		
Gambling risk (IX)	0.022	0.992	0.033	-0.011	0.001	0.071	0.014	0.074	-0.077	0.159	1	
Age (X)	20.622	1.834	-0.009	-0.024	-0.030	-0.017	-0.026	0.023	0.064	-0.071	-0.033	1
Gender	0.558	0.497	0.033	-0.041	-0.058	0.045	-0.045	0.112	-0.022	0.020	0.070	0.104

Table 6: Comparison of Actual Versus Optimal Contract Parameters and Effort Decisions

	Bonus					Penalty				
	Optimal	Canada	China	South Korea	Total	Optimal	Canada	China	South Korea	Total
Demanded effort	4	5.589***	6.300***	5.957***	5.949***	4	4.858***	4.906***	5.494***	5.089***
Fixed fee	1	13.239***	18.472***	14.319***	15.289***	5	16.442***	15.622***	18.363***	16.830***
Incentive	12	17.467***	16.750***	18.743***	17.711***	12	8.642***	9.094***	9.368***	9.034***
Effort	4	2.917***	4.061	3.195***	3.381***	4	2.474***	2.467***	2.653***	2.532***

Note: ***$p<0.01$, **$p<0.05$, *$p<0.1$; we used a simple t-test.

Table 7: Regression Results - Odds Ratios for Acceptance, Effort, and Shirking

	Bonus						Penalty					
	Accept		Effort		Shirk		Accept		Effort		Shirk	
Model	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
China	0.663	0.590	2.117***	2.674***	0.404***	0.323***	1.481	1.554	1.019	0.995	0.835	0.796
	(0.234)	(0.222)	(0.470)	(0.631)	(0.113)	(0.097)	(0.521)	(0.562)	(0.230)	(0.237)	(0.228)	(0.233)
South Korea	0.732	0.669	1.376	1.247	0.696	0.637	2.370**	2.622***	0.805	0.751	0.714	0.733
	(0.287)	(0.281)	(0.333)	(0.319)	(0.204)	(0.199)	(0.822)	(0.947)	(0.189)	(0.185)	(0.203)	(0.216)
Fixed fee	1.165***	1.170***	1.047***	1.048***	0.936***	0.939***	1.314***	1.338***	1.064***	1.058***	0.935***	0.940***
	(0.028)	(0.029)	(0.010)	(0.011)	(0.012)	(0.013)	(0.041)	(0.044)	(0.013)	(0.014)	(0.013)	(0.014)
Demanded effort	0.802**	0.803**	1.063	1.066	1.924***	1.964***	0.640***	0.598***	0.922	0.96	2.047***	2.031***
	(0.079)	(0.080)	(0.060)	(0.061)	(0.163)	(0.170)	(0.061)	(0.060)	(0.064)	(0.069)	(0.198)	(0.198)
Incentive	1.012	1.007	1.035**	1.058***	0.941***	0.925***	0.835***	0.826***	1.079***	1.088***	0.904***	0.895***
	(0.026)	(0.028)	(0.018)	(0.020)	(0.021)	(0.022)	(0.034)	(0.034)	(0.027)	(0.027)	(0.026)	(0.027)
Time trend	1.105**	1.113**	0.922***	0.917***	1.105***	1.111***	1.170***	1.165***	0.997	0.999	0.964	0.964
	(0.054)	(0.056)	(0.026)	(0.027)	(0.040)	(0.041)	(0.054)	(0.055)	(0.030)	(0.031)	(0.036)	(0.036)
Gender	0.749	0.590*	0.650**	0.554***	1.268	1.540*	0.707	0.505**	1.164	1.329	0.929	0.727
	(0.211)	(0.181)	(0.115)	(0.104)	(0.269)	(0.352)	(0.215)	(0.171)	(0.228)	(0.280)	(0.220)	(0.185)
Age	1.154	1.190*	0.982	1.018	0.924	0.909	0.821***	0.827***	1.116**	1.113**	0.880**	0.879**
	(0.114)	(0.124)	(0.060)	(0.064)	(0.067)	(0.068)	(0.058)	(0.060)	(0.056)	(0.057)	(0.056)	(0.057)
Loss aversion		0.621***		0.885		1.121		0.959		0.970		0.977
		(0.102)		(0.098)		(0.168)		(0.076)		(0.047)		(0.058)
Inequity aversion		0.776		2.443***		0.498*		0.289***		2.093***		0.446**
		(0.415)		(0.795)		(0.204)		(0.119)		(0.549)		(0.159)
Investment risk		1.064		0.686***		1.173		0.991		1.083		1.097
		(0.137)		(0.065)		(0.137)		(0.158)		(0.124)		(0.151)
Gambling risk		0.846		0.822**		1.494***		1.214		0.883		1.239*
		(0.123)		(0.081)		(0.176)		(0.162)		(0.079)		(0.137)
Observations	570	570	496	496	496	496	560	560	447	447	447	447
Log likelihood	-185.096	-178.351	-901.757	-883.07	-293.592	-284.059	-196.718	-191.163	-768.433	-763.109	-263.92	-258.871
McFadden's Pseudo R2	0.159	0.190	0.039	0.059	0.138	0.166	0.301	0.321	0.035	0.041	0.148	0.164
Correctly classified	87.37%	87.72%			67.14%	69.96%	84.82%	85.89%			67.11%	69.35%

Notes: ***p<0.01, **p<0.05, *p<0.1; standard errors in parentheses. Models I, III, V, VII, IX and XI are the main models, and Models II, IV, VI, VIII, X and XII are the models for the robustness checks discussed in Section 5.4. Fifty-seven dyads completed the bonus treatment, and 56 dyads completed the penalty treatment. Each participant played 10 rounds with random rematching; hence, the total sample sizes of 570 observations and 560 observations were used to analyze the acceptance rates in the bonus and penalty treatments,

respectively. In analyzing the level of effort and the rate of shirking, we use only those observations in which the supplier accepted an offer. Therefore, only the contracts that were accepted were used to analyze the effort level and shirking rate: Bonus contract, 496 observations (74 contracts were not accepted) and penalty contract, 447 observations (113 contracts were not accepted). Due to the binary nature of the variables, we used logistic regression for the acceptance rate and shirking models. For the model of levels of effort, we used ordered logistic analysis because this variable has an ordinal scale. Constants are not reported but are available upon request. For each dependent variable, we ran a separate model that included interaction terms; the interaction terms were not significant and therefore are dropped from the main models presented. Classification tables that report the percentages of correctly classified observations are only available for logit models. Therefore, we have reported only the percentage of correctly classified observations for acceptance and shirking models that use a binary dependent variable.

Table 8: Marginal Means for Main Models in Table 7

	Bonus (Models I, III, and V)			Difference		Penalty (Models VII, IX, and XI)			Difference	
	Canada	China	South Korea	(Canada - China)	(Canada - South Korea)	Canada	China	South Korea	(Canada - China)	(Canada - South Korea)
Accept	0.892**	0.853**	0.863**	-0.039	-0.029	0.746**	0.793**	0.843**	0.047	0.097**
Effort 1	0.412**	0.248**	0.337**	-0.154***	-0.069	0.364**	0.359**	0.415**	-0.004	0.047
2	0.040***	0.031***	0.037***	-0.006**	-0.002	0.040***	0.040***	0.042***	0.000	0.001
3	0.048***	0.041***	0.047***	-0.006**	-0.001	0.124**	0.124**	0.125**	0.000	0.000
4	0.091**	0.085**	0.091**	-0.004	0.000	0.158**	0.159**	0.149**	0.000	-0.007
5	0.162**	0.185**	0.178**	0.017**	0.011	0.170**	0.172**	0.150**	0.001	-0.016
6	0.127**	0.186**	0.153**	0.047***	0.022	0.067**	0.068**	0.057**	0.001	-0.010
7	0.078**	0.139**	0.101**	0.057***	0.022	0.038***	0.039***	0.031***	0.001	-0.001
8	0.027***	0.053**	0.036***	0.029***	0.010	0.015***	0.015***	0.012***	0.000	-0.003
9	0.007***	0.014***	0.010***	0.009**	0.003	0.008***	0.008***	0.007***	0.000	-0.002
10	0.008***	0.018***	0.012***	0.011**	0.004	0.015***	0.015***	0.012***	0.000	-0.004
Shirk	0.531**	0.347**	0.456**	-0.184***	-0.075	0.524**	0.487**	0.456**	-0.037	-0.068

Notes: ***p<0.01, **p<0.05, *p<0.1. For level of effort, we calculate the probabilities that participants make in each category of effort because we vary the country variable and hold the other variables at their means. China provides the highest effort with a higher probability under the bonus contract corresponding to a significantly positive coefficient value of China in the model.

Table 9: Regressions Results – Odds Ratios for Acceptance, Effort, and Shirking by Country

Model	Canada			China			South Korea		
	Accept (i)	**Effort (ii)**	**Shirk (iii)**	**Accept (iv)**	**Effort (v)**	**Shirk (vi)**	**Accept (vii)**	**Effort (viii)**	**Shirk (ix)**
Bonus treatment	0.473	2.886*	0.209**	0.303	2.099	0.27	1.514	0.676	1.518
	(0.458)	(1.757)	(0.162)	(0.353)	(1.545)	(0.252)	(2.027)	(0.472)	(1.266)
Fixed fee	1.205***	1.062***	0.939***	1.209***	1.081***	0.906***	1.346***	1.026*	0.958***
	(0.035)	(0.015)	(0.016)	(0.039)	(0.014)	(0.019)	(0.061)	(0.014)	(0.014)
Demanded effort	0.831*	1.072	1.980***	0.629***	1.051	2.333***	0.636***	0.98	1.776***
	(0.087)	(0.083)	(0.221)	(0.082)	(0.083)	(0.311)	(0.087)	(0.071)	(0.174)
Incentive	0.854***	1.051	0.888**	0.825***	1.075*	0.864***	0.866	1.043	0.962
	(0.052)	(0.046)	(0.049)	(0.056)	(0.046)	(0.045)	(0.076)	(0.044)	(0.047)
Bonus treatment * Incentive	1.225***	0.937	1.114*	1.221**	0.995	1.049	1.138	1.013	0.973
	(0.091)	(0.048)	(0.072)	(0.104)	(0.054)	(0.070)	(0.113)	(0.051)	(0.059)
Time trend	1.049	0.967	0.991	1.135**	0.916**	1.115**	1.294***	0.955	1.038
	(0.058)	(0.035)	(0.044)	(0.067)	(0.034)	(0.054)	(0.085)	(0.035)	(0.045)
Gender	0.900	1.148	0.902	0.647	0.8	1.135	0.542	0.672*	1.196
	(0.296)	(0.270)	(0.251)	(0.229)	(0.173)	(0.308)	(0.239)	(0.162)	(0.347)
Age	0.920	1.011	0.867*	0.696*	1.485***	0.737*	0.994	1.083	0.933
	(0.077)	(0.061)	(0.063)	(0.139)	(0.227)	(0.136)	(0.086)	(0.063)	(0.066)
Observations	370	304	304	360	301	301	400	338	338
Log likelihood	-136.906	-524.84	-180.786	-125.712	-528.354	-167.253	-116.884	-607.405	-206.442
McFadden's Pseudo R2	0.211	0.032	0.142	0.217	0.092	0.189	0.323	0.014	0.115
Correctly classified	86.49%		69.41%	87.22%		70.10%	88.00%		64.50%

Notes: ***p<0.01, **p<0.05, *p<0.1; standard errors in parentheses. The sample size is 370 observations in Canada (37 suppliers), 360 observations in China (36 suppliers), and 400 observations (40 suppliers) in South Korea. Again, we use only accepted contracts to analyze effort level and shirking rate: Canada 304 observations (66 contracts were not accepted), China 301 observations (59 contracts were not accepted), and South Korea 338 observations (62 contracts were not accepted). Because of the binary nature of the variables, we used logistic regression for the acceptance rate and shirking models. For the model of levels of effort, we use ordered logistic analysis because this variable has an ordinal scale. Constants are not reported but are available upon request. Classification tables that report the percentages of correctly classified observations are available only for logit models. Therefore, we have reported only the percentage of correctly classified observations for acceptance and shirking models that use a binary dependent variable.

Table 10: Marginal Means for Models in Table 9

		Canada			China			South Korea		
		Bonus	**Penalty**	**(Bonus-Penalty)**	**Bonus**	**Penalty**	**(Bonus-Penalty)**	**Bonus**	**Penalty**	**(Bonus-Penalty)**
Accept		0.873**	0.629**	0.244***	0.874**	0.689**	0.184***	0.908**	0.676*	0.232***
Effort	1	0.320**	0.365*	-0.045	0.238**	0.356**	-0.118**	0.454**	0.406*	0.047
	2	0.047**	0.049**	-0.002	0.021***	0.025***	-0.004	0.034**	0.034**	0.000
	3	0.062**	0.063**	-0.001	0.087**	0.098**	-0.012	0.072**	0.072**	-0.001
	4	0.183**	0.179**	0.005	0.081**	0.084**	-0.003	0.076**	0.078**	-0.003
	5	0.175**	0.161**	0.014	0.224**	0.200**	0.024*	0.116**	0.125**	-0.008
	6	0.114**	0.100**	0.014	0.160**	0.119**	0.041**	0.095**	0.106**	-0.011
	7	0.065**	0.056**	0.010	0.093**	0.061**	0.032**	0.090**	0.104**	-0.014
	8	0.021***	0.017***	0.003	0.048**	0.029**	0.019*	0.033**	0.039**	-0.006
	9	0.008***	0.007***	0.001	0.021***	0.012***	0.009	0.008***	0.010***	-0.002
	10	0.004***	0.003***	0.001	0.027**	0.015***	0.012*	0.022***	0.026**	-0.004
Shirk		0.436	0.472	-0.036	0.353	0.483	-0.130*	0.477	0.473	0.003

Note: ***p<0.01, **p<0.05, *p<0.1.

5.2. Effort in Accepted Contracts

Next, we looked at the effect of different cultural associations on the level of effort put forth in accepted contracts. For this analysis, we use an ordered logistic regression because of the ordinal scale of the dependent variable for level of effort. Our first hypothesis (H2a) is that individuals from collectivist cultures will exert more effort under bonus contracts than those from individualist cultures. Our results (Table 7, Model III) show that participants from the collectivist culture of China exerted more effort in the bonus treatment than those from the individualist culture of Canada (p<0.01). For Chinese suppliers, the odds of effort level > k versus effort level ≤ k were 2.1 times greater than for Canadian suppliers under the same conditions. However, the same relationship was only directionally true for South Korea, also a collectivist culture. The results in Table 8 on marginal means further support this finding. As can be seen, the difference in marginal means at the various levels of effort is significant in all but one instance (effort level 4) for Canada with China, but it is not significant at any level for South Korea. Thus, H2a is only partially supported. In H2b, we posited that individuals from collectivist cultures exert less effort under penalty contracts than those from individualist cultures. However, our results (Table 7, Model IX) show no differences for penalty contracts, thus H2b is not supported.[2]

Again, several of our control variables influenced the level of effort provided by the supplier. In both the bonus and the penalty treatments, we find a statistically significant and positive effect from a fixed fee on effort levels (p<0.01). As the fixed payment increases, participants are generally willing to exert more effort. Based on the results by country shown in Table 8, this effect occurs across all three countries. Across both incentive treatments, we find significant effects of the amount of incentive (i.e., bonus and penalty) offered on the suppliers' level of effort (p<0.05). Stiffer penalties and generous bonuses result in

---

[2] In H2, we assume that a supplier's risk preference is the only driver of his effort level. However, one may suspect that it is also driven by social preferences because a supplier can punish a buyer about an unfair offer by choosing a low effort level. Therefore, we looked at whether a supplier chooses an effort level depending on how unfair an offered contract is. More specifically, we calculated a difference between the expected profits of a buyer and a supplier if a supplier chooses an effort level demanded by a buyer (a measure for the degree of unfairness) and tested whether a supplier's effort level depends on this difference in expected profits. We did not find any clear pattern that a low level of effort level is associated with unfair offers, and social preferences do not seem to play an important role in predicting a supplier's effort level.

higher levels of effort. We also observe both a time trend ($p<0.01$) and a gender effect ($p<0.05$) in the bonus treatment. Suppliers in the bonus treatment tended to provide lower levels of effort over time, and males are observed to exert less effort than females. Finally, we observe that as age increases, the effort level increases only in the penalty treatment ($p<0.05$).

Results of the within-country analysis reported in Table 10 indicate that suppliers in China exert higher effort with significantly higher probability under the bonus contract (compared with the penalty contract). No such treatment effect is observed in Canada and South Korea.

5.3. Shirking in Accepted Contracts

Lastly, we evaluate the possibility of suppliers' shirking on contracts. In models V and XI in Table 7, we report the results of the logistic regression analysis for the dependent variable of shirking. We hypothesized (H3) that with both the bonus and penalty treatments, individuals from collectivist cultures would shirk less than those from individualist cultures. Our findings show support for this in only one instance; in the bonus treatment, we find that participants from the collectivist culture of China shirked significantly less than those from the individualist culture of Canada ($p<0.01$). All things being equal, Chinese participants shirked about 0.4 times as often as their Canadian counterparts. We find this only directionally true for those from the collectivist culture of South Korea. Moreover, these relationships are also only directionally supported for the penalty treatment. Consequently, support for H3 is uncertain.

Several of the control variables also contributed to shirking behavior in our experiments. We find that as the fixed fee increases, the rate of shirking decreases ($p<0.01$). As suppliers receive higher assured payments, they are likelier to comply with the contract in both the bonus and penalty settings. These findings hold across all three countries, as shown in the regressions in Table 8. We also observe that as the level of effort demanded by buyers increases, the likelihood of shirking by the suppliers increases ($p<0.01$). Again, this finding holds across all country level regressions as well. Additionally, the amount of incentive offered has a deterrent effect on shirking across both treatments. As the potential bonus or penalty increases, the likelihood of the supplier's shirking is reduced significantly ($p<0.01$). In evaluating the country analysis results, we see that this effect is comparable across bonuses and penalties but that

participants in Canada and China, not South Korea, are driving the effect. We again observe a time trend, but only in the bonus treatment ($p<0.01$) in which those suppliers are likelier to shirk in later rounds of the game. Also, age has a significant effect in the penalty treatment ($p<0.05$) in that as age increases, the rate of shirking decreases. Overall, our regression by treatment (Table 7) correctly classified about 67% of the observations correctly in both the bonus and the penalty treatments. Similar levels of fit can be observed for the regression by country (Table 9).

Further analysis of the marginal means in shirking rates within each country (Table 10) shows that suppliers in China shirk less under bonus contracts than they do in penalty contracts ($p<.10$). In the other two countries, we observe no discernible difference in shirking between the bonus and penalty treatments.

5.4. Robustness Checks

Other measures are potentially related to the supplier's decision. Loss aversion and inequity aversion (fairness) have been shown to explain the effectiveness of incentive contracts (Fehr et al., 2007; Hannan et al., 2005). In addition, our experimental set-up involves a risky decision for suppliers: They may receive a probabilistic bonus or pay a probabilistic penalty. Consequently, their attitude toward risk may explain their level of effort and their decisions about shirking. As robustness checks, we introduce additional controls for *individual* differences in loss aversion, inequity aversion, and risk attitude, and show that the observed main effects are indeed because of broader, cultural-based differences and not the result of individual differences.

We derive participants' comparative *loss aversion* score based on a choice game experiment and measure by Abdellaoui et al. (2008) and Zhang et al. (2016).[3] In addition, we use a survey scale developed by Weber et al. (2002) to collect information about participants' levels of *investment* and *gambling risk attitude* (on a 7-point Likert scale). We used confirmatory factor analysis with Varimax rotation to derive the two factors. The Cronbach's alphas are 0.731 for *investment risk* and 0.869 for

---

[3] Due to extreme outliers, we used a 1% winsorization of this variable.

gambling risk, indicating good reliability of the constructs. Both alphas are above the suggested cut-off value of 0.7 for existing scales (Nunnally and Bernstein, 1994).

Finally, during the experiment, participants were asked after every round to indicate on a 7-point Likert scale how fair they perceived the contract to be.[4] We derive a participant's average level of inequity aversion by examining the observed inequity between an individual's own payoff and their partner's payoff in each period p, $|own\ profit_p - partner's\ profit_p|$, relative to the fairness score given by the participant in the same period, $fairness\ score_p$; We then took the average of this over ten decision rounds: $inequity\ aversion = \sum_{p=1}^{10} \frac{|own\ profit_p - partner's\ profit_p|}{fairness\ score_p}$. For the same level of inequity in profits, if A gives a higher fairness score than B, we say A is less averse to inequity. Therefore, a lower value of our inequity aversion measure indicates a lower level of inequity aversion. Note that this formulation assumes that participants present disutility both when they earn less than their partners (disadvantageous inequity) and when they earn more than their partners (advantageous inequity) (Fehr and Schmidt, 1999). In other studies, fairness considerations are incorporated in the utility model so that people are motivated by both their pecuniary payoff and their relative payoff standing (see, for example, Bolton and Ockenfels (2000) and De Bruyn and Bolton (2008) in experimental games and Niederhoff and Kouvelis (2016) in the supply chain context for modeling fairness).

The regression results from these robustness checks are reported in Table 7 (models II, IV, and VI for the bonus treatment and models VIII, X, and XII for the penalty treatment). It is important to note that our main findings about cultural effects on acceptance rate, effort level, and shirking rate are robust even when including measures for individual-level loss aversion, inequity aversion, and risk attitudes as controls, i.e., the direction and significance level of the country coefficients are similar to our original

---

[4] We recognize the possibility of priming for fairness as a result of this design choice. To specifically test for priming effects in our research, we compared the level of acceptance in the first period to the last period in each country and in each treatment (bonus vs. penalty). By the last period, only those participants in China and in the bonus treatment were more likely to accept an offered contract. This finding may be because of priming or simply reflect the results of experience as participants became better able to assess what contracts were attractive to their counterparts. No such difference was observed in other instances. Further, we find no statistically significant difference in the actual effort levels or the level of shirking across these two periods.

models. The model fit improved as can be seen in the log likelihood ratios, all of which increase in value by including these controls. Some additional findings emerge from the new model (significant at the 5% level). We find that an individual's loss aversion score is negatively associated with acceptance rates in the bonus treatment, in that participants who are more loss averse are less likely to accept a bonus contract. Furthermore, greater inequity aversion in the penalty treatment results in fewer accepted contracts. In both the bonus and the penalty treatments, inequity aversion has a significant effect on the level of effort; participants more averse to inequity put forth more effort. In addition, in the penalty treatment, inequity aversion has a significant effect on shirking such that participants more averse to inequity are less likely to shirk. Finally, we find in the bonus treatment that individuals' risk attitudes are closely related to effort levels and shirking decisions; suppliers who are risk-seeking tend to offer lower levels of effort and shirk more.

Finally, given the use of repeated observations for each subject as each participant played 10 rounds, we created a fully balanced panel data set. This allows us to control for correlations between participants' behavior over the 10 periods. We find the results to be consistent to the findings in Table 7.

6. Discussion and Implications

Scholars have recently developed an interest in better understanding how and why culture matters in supply chain exchanges (Cannon et al., 2010; Metters et al., 2010). Our research contributes to this discourse by theoretically and empirically evaluating the interplay of three behavioral considerations (fairness, risk, and opportunism) with national culture on the effectiveness of incentive contracts. Previous studies have realized the impact of these behavioral aspects on incentive contract effectiveness but in isolation from cultural considerations. This is a key oversight because these behaviors do not have universal effects across cultures (Doney et al., 1998). This means that previous research findings can benefit from a more nuanced understanding of how culture matters in this context. Our results show that contract effectiveness (i.e., contract acceptance, supplier effort, and supplier shirking) varies based on this interplay.

The findings from our study culminate in both theoretical and practical contributions. First, our research offers a distinct contribution about notions of fairness where incentive contracts are concerned. Prior studies have noted an aversion to inequity that drives higher acceptance rates for bonus contracts (Fehr et al., 2007; Hannan et al., 2005). As different cultures value fairness differently (Henrich et al., 2005), we argue in this research that how this difference affects the acceptance levels of contracts is more nuanced than the previous literature reflects. Acceptance rates for the bonus treatment are not significantly different in any of the three nationalities and average 85-89%. However, we find substantive cultural differences in acceptance rates for penalty contracts. In the penalty treatment, supplier acceptance rates ranged from a low of 74.6% for Canadian participants to a high of 84.3% for South Korean participants – a difference that is statistically significant. Moreover, odds ratio results indicate that South Koreans were about 2.4 times as likely to accept a penalty contract as their Canadian counterparts. These results are consistent with the theoretical rationale underlying the cultural dimensions of power distance, which reflects the perceived procedural fairness of the contract offered, and with humane orientation, which reflects societal emphasis and encouragement of fair action (House et al., 2004). Our research complements a burgeoning stream of work focused on the impact of fairness in supply chain contracting (e.g., Ho et al., 2014; Katok and Pavlov, 2013; Katok and Wu, 2009), by evaluating the interplay of fairness and national culture on the effectiveness of incentives. Managers should be especially concerned with the punishment of unfair actions because the cultural differences we observed were most notable in the penalty incentive contracts. Cultures associated with low power distance and high humane orientation are likely to be more sensitive to issues of fairness. Members of these cultures may have exceptionally low rates of acceptance of penalty contracts.

We also evaluated the interplay of risk taking, opportunistic behavior, and national culture on a supplier's level of effort and shirking. Previous research has expounded a "cushion hypothesis" (Hsee and Weber, 1999), which we leverage to predict that individuals from collectivist cultures would exert more (less) effort under bonus (penalty) contracts than those from individualist cultures. Our findings only partially support this notion. In the penalty treatment, once a contract was accepted, suppliers across all

three nationalities submitted very similar levels of effort. However, in the bonus treatment, collectivist Chinese suppliers put forth significantly higher levels of effort than individualist Canadian suppliers. For Chinese suppliers, all things being equal, the odds of effort level > k versus effort level ≤ k were 2.1 times greater than for Canadian suppliers. The differential observed between effort levels in Canada and China fits with our cultural supposition. However, the relationship holds only directionally when comparing collectivist South Korean suppliers and individualist Canadian suppliers. This suggests that another mechanism beyond individualism must be at work in considering the level of effort a supplier chooses to dedicate to a contract. One possible explanation is that South Korean society is more *assertive* than Canadian and Chinese societies (House et al., 2004). Many of the attributes associated with the dimension of assertiveness run counter to those of a collectivist culture; assertive societies value competition, success, and performance. It may be that assertive tendencies balance those associated with collectivism in this environment. Indeed, prior research argues that assertiveness is positively associated with opportunism (Doney et al., 1998); yet, our South Korean participants were no likelier to shirk than Canadian participants in either treatment. This seems to suggest that South Koreans' cultural assertiveness is indeed balanced – but not supplanted – by collectivist tendencies. Moreover, low assertiveness, in conjunction with collectivism, may account for the exceptionally low shirking behavior we observe for our Chinese participants in the bonus treatment. The combination of collectivism and low assertiveness would almost certainly reinforce behavior that is cooperative, harmonious, and high in integrity (House et al., 2004). However, research on assertiveness as a cultural dimension is scant. Our findings suggest a need for further investigation of how assertiveness manifests behaviorally as well as the need for a better understanding of its interactive effect with other cultural dimensions. Our results both complement and extend those of Handley and Angst (2015) by further investigating how culture influences opportunistic action in supply chain exchanges and also by extending this evaluation to incorporate a broader set of cultural dimensions.

Finally, our within-country analysis on contract effectiveness revealed some interesting empirical findings that are relevant and important for managers. Overall, we show that bonus contracts are accepted

more than penalty contracts in all three countries. Especially given that incentive contracts can be designed to be economically equivalent, we suggest managers working within any of these cultures consider adopting bonus framing, if for no other reason than to facilitate fast and efficient acceptance from their suppliers. Once contracts are accepted, bonus contracts are more effective in China because of higher levels of effort and fewer instances of shirking. We further show that for accepted contracts, the effectiveness of bonus and penalty contracts is nearly indistinguishable in Canada and South Korea. As such, the benefits of a bonus incentive in contract performance (supplier effort and non-shirking) may be limited to a small and specific cultural profile. In such circumstances, other concerns, such as financial or enforceability issues, may prove to be stronger decisional drivers.

7. Limitations and Conclusions

This research is subject to certain limitations, many of which also open the door for future research opportunities. First, our assessments of culture are formed at the national level, leveraging cultural frameworks and dimensions prevalent in the literature. We recognize that studies incorporating cultural measures at the level of the individual, as opposed to a national level, may be a useful extension to this research. Additionally, our findings suggest that additional dimensions of national culture are likely influencing the decision-making behaviors observed in this context. Future research parsing out further potential dimensional peculiarities could help to better align managerial decisions across cultures beyond those studied here. Second, it could prove useful to evaluate the behaviors of individuals subjected to both types of incentive contracts because those in our study performed under just one treatment: bonus *or* penalty. Exposure to both contract types would likely further influence behavior. Our expectation is that this type of exposure would serve to exacerbate the relationships found in our research, and also perhaps detect new ones.

A final limitation to our work involves our decision to focus exclusively on an intra-cultural comparison of contract effectiveness, a decision made because of the novelty and depth of theorizing. We strongly emphasize that the results we find in this research do not necessarily transfer to inter-cultural

supply chain exchanges. Several of the cultural dimensions under consideration in this work, including collectivism and humane orientation, are known to have in-group versus out-group orientations (Schlösser et al., 2013; Triandis, 1983). This means how individuals behave with other individuals from the same national culture (our intra-cultural supply chains) may vary dramatically from how they would interact with those from another culture (an inter-cultural perspective). This was found to be the case in Özer et al. (2014) evaluation of trust and trustworthiness across U.S. and Chinese cultures. For example, the Chinese were more trusting of their U.S. supply chain counterparts than they were of Chinese ones. Similarly, cultural distance was also determined to be critical in understanding how justice-based governance practices impact the experience of conflict in buyer-supplier relationships (Srinivasan et al., 2018). It would be informative to extend our line of research to those involving buyer and supplier groups of different cultures in an inter-cultural examination (Zaheer and Kamal, 2011), which would require a substantially different line of theorizing beyond the scope of the current effort.

Notwithstanding, the contributions made with this research are both timely and relevant well beyond a foundational piece for future inter-cultural work. In a time of increased reshoring (Ellram et al., 2013; Foerstl et al., 2016; Gray et al., 2013) caused in part by technological developments that have changed the financial landscape and by rising nationalism (Lehmacher, 2017) among many countries because of their concerns over security and competitiveness, this work squarely places emphasis on managing these newly shortened supply chains.

Our work contributes to the supply chain contracting literature by incorporating an in-depth evaluation at the intersection of national culture and behavioral perceptions and their impact on incentive contract effectiveness. In conclusion, there are ample opportunities for further research evaluating the effectiveness of incentive contracts through a cultural lens.

References

Abdellaoui, M., Bleichrodt, H., l'Haridon, O., 2008. A tractable method to measure utility and loss aversion under prospect theory. Journal of Risk and Uncertainty 36, 245-266

Alchian, A.A., Demsetz, H., 1972. Production, Information Costs, and Economic Organization. The American Economic Review 62, 777-795

Aron, R., Clemons, E.K., Reddi, S., 2005. Just Right Outsourcing: Understanding and Managing Risk. Journal of Management Information Systems 22, 37-55

Berry, J.W., 1997. Immigration, Acculturation, and Adaptation. Applied Psychology: An International Review 46, 5-68

Berry, J.W., Poortinga, Y.H., Segall, M.H., Dasen, P.R., 1992. Cross-cultural psychology: Research and applications. Cambridge University Press, New York

Bigoni, M., Fridolfsson, S.-O., Le Coq, C., Spagnolo, G., 2012. Fines, leniency, and rewards in antitrust. The RAND Journal of Economics 43, 368-390

Bloom, M., Milkovich, G.T., 1998. Relationships among risk, incentive pay, and organizational performance. Academy of Management Journal 41, 283-297

Bolton, G.E., Ockenfels, A., 2000. ERC: A theory of equity, reciprocity, and competition. American Economic Review 90, 166-193

Bolton, G.E., Ockenfels, A., Thonemann, U.W., 2012. Managers and students as newsvendors. Management Science 58, 2225-2233

Breuer, W., Riesener, M., Salzmann, A.J., 2014. Risk aversion vs. individualism: What drives risk taking in household finance? The European Journal of Finance 20, 446-462

Brockner, J., Ackerman, G., Greenberg, J., Gelfand, M.J., Francesco, A.M., Chen, Z.X., Leung, K., Bierbrauer, G., Gomez, C., Kirkman, B.L., Shapiro, D., 2001. Culture and Procedural Justice: The Influence of Power Distance on Reactions to Voice. Journal of Experimental Social Psychology 37, 300-315

Buchan, N., Croson, R., 2004. The boundaries of trust: Own and others' actions in the US and China. Journal of Economic Behavior & Organization 55, 485-504

Cannon, J.P., Doney, P.M., Mullen, M.R., Petersen, K.J., 2010. Building long-term orientation in buyer-supplier relationships: The moderating role of culture. Journal of Operations Management 28, 506-521

Cavusgil, S.T., Deligonul, S., Zhang, C., 2004. Curbing foreign distributor opportunism: An examination of trust, contracts, and the legal environment in international channel relationships. Journal of International Marketing 12, 7-27

De Bruyn, A., Bolton, G.E., 2008. Estimating the influence of fairness on bargaining behavior. Management Science 54, 1774-1791

Doney, P.M., Cannon, J.P., Mullen, M.R., 1998. Understanding the Influence of National Culture on the Development of Trust. Academy of Management Journal 23, 601-620

Eckerd, S., Boyer, K.K., Qi, Y., Eckerd, A., Hill, J.A., 2016. Supply chain psychological contract breach: An experimental study across national cultures. Journal of Supply Chain Management 52, 68-82

Eisenhardt, K.M., 1989. Building Theories from Case Study Research. Academy of Management Review 14, 532-550

Ellram, L.M., Tate, W.L., Billington, C., 2004. Understanding and managing the services supply chain. Journal of Supply Chain Management 40, 17-32

Ellram, L.M., Tate, W.L., Petersen, K.J., 2013. Offshoring and reshoring: An update on the manufacturing location decision. Journal of Supply Chain Management 49, 14-22

Fehr, E., Falk, A., 2002. Psychological foundations of incentives. European Economic Review 46, 687-724

Fehr, E., Gächter, S., 2000. Cooperation and Punishment in Public Goods Experiments. The American Economic Review 90, 980-994

Fehr, E., Gächter, S., Kirchsteiger, G., 1997. Reciprocity as a contract enforcement device: Experimental evidence. Econometrica 65, 833-860

Fehr, E., Klein, A., Schmidt, K.M., 2007. Fairness and Contract Design. Econometrica 75, 121-154

Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. The Quarterly Journal of Economics 114, 817-868

Fehr, E., Schmidt, K.M., 2004. Fairness and Incentives in a Multi-task Principal–Agent Model. The Scandinavian Journal of Economics 106, 453-474

Fehr, E., Schmidt, K.M., 2007. Adding a stick to the carrot? The interaction of bonuses and fines. American Economic Review 97, 177-181

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made econometric experiments. Experimental Economics 10, 171-178

Foerstl, K., Kirchoff, J.F., Bals, L., 2016. Reshoring and insourcing: Drivers and future research directions. International Journal of Physical Distribution and Logistics Management 46, 492-515

Ghoshal, S., Moran, P., 1996. Bad for Practice: A Critique of the Transaction Cost Theory. Academy of Management Review 21, 13-47

Giacobbe-Miller, J.K., Miller, D.J., Zhang, W., Victorov, V.I., 2003. Country and organizational-level adaptation to foreign workplace ideologies: A comparative study of distributive justice values in China, Russia and the United States. Journal of International Business Studies 34, 389-406

Gray, J.V., Massimino, B., 2014. The effect of language differences and national culture on operational process compliance. Production and Operations Management 23, 1042-1056

Gray, J.V., Skowronski, K., Esenduran, G., Rungtusanatham, M.J., 2013. The reshoring phenomenon: What supply chain academics ought to know and should do. Journal of Supply Chain Management 49, 27-33

Handley, S.M., Angst, C.M., 2015. The Impact of Culture on the Relationship between Governance and Opportunism in Outsourcing Relationship. Strategic Management Journal 36, 1412-1434

Handley, S.M., Benton, W.C.J., 2012. The Influence of Exchange and Power on Opportunism in Outsourcing Relationships. Journal of Operations Management 30, 55-68

Hannan, R.L., Hoffman, V.B., Moser, D.V., 2005. Bonus versus Penalty: Does Contract Frame affect employee effort? Experimental Business Research 2, 151-169

Heine, S.J., Lehman, D.R., Markus, H.R., Kitayama, S., 1999. Is there a universal need for positive self regard? Psychological Review 106, 766-794

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., Alvard, M., Barr, A., Ensminger, J., Smith Henrich, N., Hill, K., Gil-White, F., Gurven, M., Marlowe, F.W., Patton, J.Q., Tracer, D., 2005. "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. Behavioral and Brain Sciences 28, 795-815

Ho, T.-H., Su, X., Qu, Y., 2014. Distributional and Peer-Induced Fairness in Supply Chain Contract Design. Production and Operations Management 23, 161-175

Hofstede, G., 1980. Culture and Organizations. International Studies of Management & Organization 10, 15-41

Hofstede, G., 1991. Culture and Organizations: Software of the Mind. McGraw-Hill, London

House, R.J., Hanges, P.J., Javidan, M., Dorfman, P.W., Gupta, V., 2004. Culture, Leadership, and Organizations - The GLOBE Study of 62 Societies. SAGE Publications, Thousand Oaks

Hsee, C.K., Weber, E.U., 1999. Cross-national differences in risk preference and lay predictions. Journal of Behavioral Decision Making 12, 165-179

Katok, E., Pavlov, V., 2013. Fairness in supply chain contracts: A laboratory study. Journal of Operations Management 31, 129-137

Katok, E., Wu, D.Y., 2009. Contracting in Supply Chains: A Laboratory Investigation. Management Science 55, 1953-1968

Ketchen, D.J.J., Hult, G.T.M., 2007. Bridging organization theory and supply management: The case of best value supply chains. Journal of Operations Management 25, 573-580

Kühberger, A., Schulte-Mecklenbeck, M., Perner, J., 1999. The effects of framing, reflection, probability, and payoff on risk preference in choice tasks. Organizational Behavior and Human Decision Processes 78, 204-231

Kvaløy, O., Olsen, T.E., 2016. Incentive provisions when contracting is costly. Economica 83

Lam, S.K., Schaubroeck, J., Aryee, S., 2002. Relationship between organizational justice and employee work outcomes: a cross-national study. Journal of Organizational Behavior 23, 1-18

Lehmacher, W., 2017. The Global Supply Chain - How Technology and Circular Thinking Transform our Future. Springer International Publishing

Lewis, G., Bajari, P., 2014. Moral hazard, incentive contracts, and risk: evidence from procurement. Review of Economic Studies 81, 1201-1228

Li, A., Cropanzano, R., 2009. Fairness at the group level: Justice climate and intraunit justice climate. Journal of Management 35, 564-599

Liu, Y., Huang, Y., Luo, Y., Zhao, Y., 2012. How does justice matter in achieving buyer–supplier relationship performance? Journal of Operations Management 30, 355-367

Loch, C.H., Wu, Y., 2008. Social preferences and supply chain performance: An experimental study. Management Science 54, 1835-1849

Luo, Y., 2007. An integrated anti-opportunism system in international exchange. Journal of International Business Studies 38, 855-877

Meisel, M.K., Ning, H., Campbell, W.K., Goodie, A.S., 2016. Narcissism, Overconfidence, and Risk Taking in US and Chinese Student Samples. Journal of Cross-Cultural Psychology 47, 385-400

Metters, R., Zhao, X., Bendoly, E., Jiang, B., Young, S., 2010. "The way that can be told of is not an unvarying way": Cultural impacts on Operations Management in Asia. Journal of Operations Management 28, 177-185

Miras-Rodriguez, M.M., Carrasco-Gallego, A., Escobar-Perez, B., 2015. Are socially responsible behaviors paid off equally? A Cross-Cultural Analysis. Corporate Social Responsibility and Environmental Management 22, 237-256

Moorman, R.H., Balkely, G.L., 1995. Individualism-collectivism as an individual difference predictor of organizational citizenship behavior. Journal of Organizational Behavior 16, 127-142

Morrison, E.W., Robinson, S.L., 1997. When employees feel betrayed: A model of how psychological contract violation develops. Academy of Management Review 22, 226-256

Narasimhan, R., Narayanan, S., Srinivasan, R., 2013. An investigation of justice in supply chain relationships and their performance impact. Journal of Operations Management 31, 236-247

Niederhoff, J.A., Kouvelis, P., 2016. Generous, spiteful, or profit maximizing suppliers in the wholesale price contract: A behavioral study. European Journal of Operational Research 253, 372-382

Nunnally, J.C., Bernstein, I.H., 1994. The assessment of reliability. Psychometric Theory 3, 248-292

Ong, L.L., 1997. Burgernomics: The Economics of the Big Mac Standard. Journal of International Money and Finance 16, 865-878

Özer, Ö., Zheng, Y., Ren, Y., 2014. Trust, Trustworthiness, and Information Sharing in Supply Chains Bridging China and the United States. Management Science 60, 2435-2460

Parks, C.D., Vu, A.D., 1994. Social dilemma behavior of individuals from highly individualist and collectivist cultures. Journal of Conflict Resolution 38, 708-718

Peng, K., Nisbett, R.E., 1999. Culture, dialectics, and reasoning about contradiction. American Psychologist 54, 741-755

Prendergrast, C., 2002. The tenuous trade-off between risk and incentives. The Journal of Political Economy 110, 1071-1102

Ribbink, D., Grimm, C.M., 2014. The impact of cultural differences on buyer-supplier negotiations: An experimental study. Journal of Operations Management 32, 114-126

Roels, G., Karmarkar, U.S., Carr, S., 2010. Contracting for collaborative services. Management Science 56, 849-863

Sanders, N.R., Locke, A., Moore, C.B., Autry, C.W., 2007. A multidimensional framework for understanding outsourcing arrangements. Journal of Supply Chain Management 43, 3-15

Schlösser, O., Frese, M., Heintze, A.-M., Al-Najjar, M., Arciszewski, T., Besevegis, E., Bishop, G.D., Bonnes, M., Clegg, C.W., Drozda-Senkowska, E., Gaborit, M., 2013. Humane orientation as a new cultural dimension of the globe project: A validation study of the globe scale and out-group humane orientation in 25 countries. Journal of Cross-Cultural Psychology 44, 535-551

Shou, Z., Zheng, X.V., Zhu, W., 2016. Contract ineffectiveness in emerging markets: An institutional theory perspective. Journal of Operations Management 46, 38-54

Simon, M., Houghton, S.M., Aquino, K., 2000. Cognitive biases, risk perception, and venture formation: How individuals decide to start companies. Journal of Business Venturing 15, 113-134

Srinivasan, R., Narayanan, S., Narasimhan, R., 2018. An investigation of justice, conflict, and moderating effects of supplier autonomy and cultural distance in buyer-supplier relationships. IEEE Transactions on Engineering Management 65, 6-20

Stevens, C.K., 2011. Questions to consider when selecting student samples. Journal of Supply Chain Management 47, 19-21

Sun, C., Edara, P., Mackley, A., 2013. Refocusing on liquidated damages in incentive/disincentive contracts. Journal of Legal Affairs and Dispute Resolution in Engineering Construction 5, 136-141

Triandis, H.C., 1983. Collectivism versus individualism: A Reconceptualization, In: Illinois, U.o. (Ed)

Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: Heuristics and biases. Science 185, 1124-1131

Tversky, A., Kahneman, D., 1986. Rational choice and the framing of decisions. The Journal of Business 59, S251-S278

Tyler, T.R., Lind, E.A., Huo, Y.J., 2000. Cultural values and authority relations: The psychology of conflict resolution across cultures. Psychology, Public Policy, and Law 6, 1138-1163

Unger, B.N., Rank, J., Gemunden, H.G., 2014. Corporate innovation culture and dimensions of project portfolio success: The moderating role of national culture. Project Management Journal 45, 38-57

Weber, E.U., Blais, A., Betz, N.E., 2002. A Domain-Specific Risk-Attitude Scale: Measuring Risk Perceptions and Risk Behaviors. Journal of Behavioral Decision Making 15, 263-290

Weber, E.U., Hsee, C.K., 1998. Cross-cultural differences in risk perception, but cross-cultural similarities in attitudes towards perceived risk. Management Science 44, 1205-1217

Weber, L., Mayer, K.J., 2011. Designing Effective Contracts: Exploring the Influence of Framing and Expectations. Academy of Management Review 36, 53-75

Williamson, O.E., 1985. The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting. Free Press, New York

Williamson, O.E., 1991. Comparative economic organization: The analysis of discrete alternatives. Administrative Science Quarterly 36, 269-296

Woods, D.L., 1999. Incentive Clauses in Contracts, In: ISM, I.f.S.M. (Ed). ISM, Institute for Supply Management

Wu, Y., Loch, C.H., Ahmad, G., 2011. Status and relationships in social dilemmas of teams. Journal of Operations Management 29, 650-662

Wuttke, D.A., Donohue, K., Siemsen, E., 2013. Initiating and Sustaining Supplier Involvement in Development Projects: Behavioral Aspects in the Contract Design,

Xue, M., Field, J.M., 2008. Service coproduction with information stickiness and incomplete contracts: Implications for consulting services design. Production and Operations Management 17, 357-372

Zaheer, A., Kamal, D.F., 2011. Creating trust in piranha-invested waters: The confluence of buyer, supplier and host country contexts. Journal of International Business Studies 42, 48-55

Zhang, Y., Donohue, K., Cui, T.H., 2016. Contract Preferences and Performance for the Loss-Averse Supplier: Buyback vs. Revenue Sharing. Management Science 62, 1734-1754

Appendix

A.1. Analysis of the Optimal Contract in the Self-Interest Model

In a first best world, the optimal contract is achieved when the buyer does not invest a fixed cost k for verification and the supplier chooses an effort level that maximizes v(e) – c(e). Below, we analyze the optimal contract design for the self-interested buyer and supplier, i.e. each player is self-interested in maximizing their expected payoffs.

A.1.1. Bonus Contract. Under a bonus contract, the buyer should choose the minimum bonus to maximize their payoff, i.e., $b = \bar{b}$. Let $\hat{e}_B$ be the highest incentive-compatible effort level so that

$$w_B - c(\underline{e}_B) \le w_B - c(\hat{e}_B) + p\bar{b}$$

Recall $c(\underline{e}_B) = 0$. Then the buyer demands the highest incentive compatible effort level $e_B^* = \hat{e}_B$ so that

$$c(e_B^*) = p\bar{b}.$$

Given the participation constraint as

$$w_B - c(\hat{e}_B) + p\bar{b} \ge \lambda,$$

the buyer should offer a fixed fee so that

$$w_B = c(e_B^*) - p\bar{b} + \lambda.$$

In equilibrium, the payoffs of the supplier and buyer are

$$M_B^S = \lambda$$

$$M_B^B = v(e_B^*) - w_B - k - p\bar{b}.$$

A.1.2. Penalty Contract. If a buyer wants to maximize their payoff, they should choose the maximum penalty, i.e., $f = \bar{f}$. Thus, the incentive compatibility constraint is

$$w_P - c(\underline{e}_P) - p\bar{f} \le w_P - c(\hat{e}_P).$$

Thus, the buyer should require the highest incentive-compatible effort level $e_P^* = \hat{e}_P$ so that

$$c(e_P^*) = p\bar{f}$$

Given the participation constraint as

$$w_P - c(\hat{e}_P) \geq \lambda,$$

which guarantees a supplier a payoff of at least λ (reservation utility) by their choice $\hat{e}_P$, the buyer should offer a fixed fee so that

$$w_P = c(e_P^*) + \lambda.$$

In equilibrium, the payoffs of the supplier and buyer are

$$M_P^S = \lambda$$

$$M_P^B = v(e_P^*) - w_P - k.$$

A.2. Economically Equivalent Contracts

Although our main analysis is to compare the effectiveness of a *given incentive contract* (bonus or penalty) across cultures, our investigation also extends to comparisons of the effectiveness of different contract types in a *given cultural setting*. To investigate the comparative effectiveness of contractual structures on performance in terms of only contract framing (bonus versus penalty), it is imperative that the penalty contract be economically equivalent to the bonus contract. Proposition 1 shows that the condition for economically equivalent contracts is $\bar{f} = \bar{b}$ for the self-interested players analyzed in 3.2:

PROPOSITION 1. If $\bar{f} = \bar{b}$, (i) $e_P^*=e_B^*$; (ii) $w_P = w_B + p\bar{b}$; (iii) $M_P^S = M_B^S = \lambda$; (iv) $M_P^B = M_B^B = v(e_P^*) - w_P - k = v(e_B^*) - w_B - k - p\bar{b}$

PROOF. To see (i), recall that $c(e_P^*) = p\bar{f}$ and $c(e_B^*) = p\bar{b}$. Since $\bar{f} = \bar{b}$, $e_P^*=e_B^*$. To see (ii), recall that $w_P = c(e_P^*)$ and $w_B = c(e_B^*) - p\bar{b}$. Since $c(e_P^*) = c(e_B^*)$ from (i), $w_P = w_B + p\bar{b}$. It then becomes trivial to show that (iii) and (iv) hold by substituting (i) and (ii) into the equilibrium payoff functions in the bonus and penalty contracts.