# *Trust, Trustworthiness, and Information Sharing in Supply Chains Bridging China and the United States*

# Trust, Trustworthiness, and Information Sharing in Supply Chains Bridging China and the U.S.

Özalp Özer  Yanchong Zheng  Yufei Ren

Jindal School of Management, The University of Texas at Dallas, Richardson, TX 75080, oozer@utdallas.edu
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02142, yanchong@mit.edu
Department of Economics, Union College, Schenectady, NY 12308, reny@union.edu

*Management Science*, 2014, 60(10): 2435-2460

Whether and how trust and trustworthiness differ between a collectivist society, e.g., China, and an individualistic one, e.g., the U.S., generate much ongoing scientific debate and bear significant practical values for managing cross-country transactions. We experimentally investigate how supply chain members' countries of origin – China versus the U.S. – affect trust, trustworthiness, and strategic information sharing behavior in a cross-country supply chain. We consider a two-tier supply chain in which the upstream supplier solicits demand forecast information from the retailer to plan production; but the retailer has an incentive to manipulate her forecast to ensure abundant supply. The levels of trust and trustworthiness in the supply chain and supplier's capability to determine the optimal production quantity affect the efficacy of forecast sharing and the resulting profits. We develop an experimental design to disentangle these three aspects and to allow for real-time interactions between geographically distant and culturally heterogeneous participants. We observe that, when there is no prospect for long-term interactions, our Chinese participants consistently exhibit lower spontaneous trust and trustworthiness than their U.S. counterparts do. We quantify the differences in trust and trustworthiness between the two countries, and the resulting impact on supply chain efficiency. We also show that Chinese individuals exhibit higher spontaneous trust towards U.S. partners than Chinese ones, primarily because they perceive that individuals from the U.S. are more trusting and trustworthy in general. This positive perception towards U.S. people is indeed consistent with the U.S. participants' behavior in forecast sharing. In addition, we quantify that a Chinese supply chain enjoys a larger efficiency gain from repeated interactions than a U.S. one does, as the prospect of building a long-term relationship successfully sustains trust and trustworthiness by Chinese partners. We advocate that companies can reinforce the positive perception of Westerners held by the Chinese population and commit to long-term relationships to encourage trust by Chinese partners. Finally, we also demonstrate that both populations exhibit similar pull-to-center bias when solving a decision problem under uncertainty (i.e., the newsvendor problem).*

*Keywords*: Trust; trustworthiness; collectivism; individualism; western stereotypes; *guanxi*; China; forecast information; behavioral economics; experimental economics

*History*: This paper was first submitted on June 25, 2011; revised and resubmitted on October 8, 2012, May 26, 2013, November 21, 2013; and accepted on December 13, 2013.

---

* The authors thank Yossi Aviv, Gerard Cachon, Rachel Croson, Warren Hausman, Elena Katok, Stephen Leider, Upender Subramanian, an anonymous reviewer from the Russell Sage Foundation, an anonymous associate editor, and three anonymous

## 1. Introduction

In 2010, China passed Japan to become the second largest economy in the world, behind the United States (Barboza 2010). The merchandise trade volume between China and the U.S. alone accounts for 3% of the total world trade volume in 2012 (World Trade Organization 2013). With China becoming the fastest-growing economy and the world's largest manufacturer, a large number of Fortune 500 companies (e.g., Apple, Ford Motors, Intel, Procter & Gamble) rely more and more on material supply or component manufacturing in China for their products. Therefore, maintaining a cooperative supply chain relationship with their Chinese partners gains increasing importance in these companies' core competency. A key issue in managing such a global supply chain is the heterogeneity in culture and institutions between China and the U.S.

Companies that overlook the distinct cultural and market characteristics in China but simply replicate their U.S. business models have encountered bitter failures in the Chinese market. A well-known example is the loss of eBay, the leading internet trading company in the U.S., to the local Chinese competitor taobao.com. The loss of eBay is often attributed to the lack of understanding in local consumers by eBay's management team. For instance, eBay did not allow online chatting between buyers and sellers because of concerns that they might close the transactions offline to circumvent fees. In contrast, taobao.com allowed such conversations to help cultivate trust among the trading parties (Barboza and Stone 2010, Wang 2010). As another example, executives from Quantum Corporation commented that the common practice of rotating management positions in the company created a big hurdle for building a trusting relationship with its Japanese partner (Hausman 2011). This is because Japanese, like Chinese, are relationship-oriented and rely heavily on long-term contact with the same individual to develop trusting relationships. Similar issues can arise even within an organization. For example, senior executives at Hitachi – Japan pointed out difficulties in establishing relationships with managers and business analysts in Hitachi's U.S. headquarters because the U.S. management team rotates positions more frequently than their counterpart in Japan (Dyer 2011). These and many other examples have highlighted the importance of understanding China's distinctive cultural and institutional characteristics in efficient operations and supply chain management.

Our goal in this paper is to determine whether and how supply chain members' countries of origin, China versus the U.S., may affect the efficiency of strategic supply chain interactions in

reviewers for their constructive feedback. The discussions during the authors' presentations at the 2012 Behavioral Operations Workshop (Washington, DC); Northwestern University; Pennsylvania State University; University of California at Los Angeles; University of Maryland at College Park; and University of Southern California were also beneficial. We are thankful to Beth Pickett and the staff in the Center for Behavioral and Experimental Economic Science Laboratory at the University of Texas at Dallas for logistics support of the experiments in the U.S. We also appreciate the generous support offered by Jian Chen, Bo Mao, Xia Xu and the staff in the Management Information System Laboratory at Tsinghua University for the experiments in China. The authors gratefully acknowledge financial support from National Science Foundation Grant 1002381 and Russell Sage Foundation Grant 98-10-10. Comments and suggestions are highly appreciated and welcomed.

information-critical transactions. To do so, we focus on the important operations problem of information sharing. Today's complex products such as computers, aircrafts, and motor vehicles are mostly built in supply chains extended across China and the U.S. In this environment, suppliers need accurate forecast information from manufacturers or retailers to plan for production. Large companies like General Motors, Procter & Gamble, and Neiman Marcus invest heavily in deploying information management systems within their global supply chains to better coordinate with their suppliers around the world (The Economist 2008, CDC White Paper 2011). However, the values of these information systems largely depend on whether information is credibly shared. Consider the example of a two-tier supply chain in which the upstream supplier solicits demand forecast information from the retailer to plan production; but the retailer has an incentive to manipulate her forecast to ensure abundant supply. Forecast manipulation in the form of reporting overoptimistic forecasts is prevalent across industries ranging from electronics, semiconductors, medical equipment, to commercial aircraft (Lee et al. 1997, Cohen et al. 2003). Earlier research has suggested that designing contracts to align pecuniary incentives (Cachon and Lariviere 2001, Özer and Wei 2006), engaging in repeated interactions (Ren et al. 2010), and the parties' inherent propensity for trust and trustworthiness (Özer et al. 2011) all help to mitigate the extent of forecast manipulation. Conversely, the lack of trust by upstream suppliers in downstream buyers' demand forecasts often limits the supply chain's ability to satisfy final market demand. For example, the Boeing Company, who works with 17,500 suppliers in more than 50 countries, commented that "it has sometimes been a job to persuade all these suppliers to invest enough to meet future demand." The company learned that an effective way to do so is to build more trust in the supply chain and be more open to share information with their suppliers (The Economist 2012). Along similar lines, Kumar (1996) advocates that trust is an important force to help sustain effective supply chain information sharing and collaboration. Özer et al. (2011) conducted the first controlled laboratory experiments to show the role of trust in ensuring effective demand forecast information sharing in a supply chain. The current paper builds on their experimental set-up to incorporate a global perspective.

Trust and its interaction with culture are important factors that influence the efficiency of economic transactions. "Virtually every commercial transaction has within itself an element of trust" (Arrow 1972, p. 357). "Culture, for purposes of economic organization, serves as a check on opportunism" (Williamson 1993, p. 476). A commonly agreed definition of trust stipulates that "trust is a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another" (Rousseau et al. 1998, p. 395). Trust has been extensively studied with the trust game (Kreps 1990) and the investment game (Berg et al. 1995) along the dimension of property rights; the trustor voluntarily passes his property rights to the trustee in hope of the trustee's reciprocity. In this paper, we study how culture affects trust

in the dimension of strategic information communication. In the forecast sharing setting, trust controls the degree to which the supplier is willing to rely on the retailer's forecast report to plan production, subjecting himself to the vulnerability of building excess inventory; and trustworthiness controls the retailer's incentive to manipulate her forecast information. Our results help to quantify the impact of culture on trust, trustworthiness, and the resulting information sharing efficiency in a concrete business context.

How trust and trustworthiness develop vary in different countries (Doney et al. 1998). To investigate these potentially different dynamics, we design our experiments to unambiguously study spontaneous trust and trustworthiness versus the temporal evolution of these factors. Spontaneous trust and trustworthiness refer to an individual's tendency to trust and to be trustworthy towards a partner with whom there is no history of social interactions. We first study a single-interaction setting to disentangle spontaneous trust from mechanisms such as reputation effects which may motivate trust in repeated interactions (Berg et al. 1995). In addition, the single-interaction set-up is also practically important. Economists have shown that the level of spontaneous trust in the general population is a strong indicator for market and economic efficiency in a society (Knack and Keefer 1997, La Porta et al. 1997, Zak and Knack 2001). Such spontaneous trust also serves as the starting point of trust evolution in a relationship. Companies (particularly the global ones) constantly consider and invest in new business relationships without prior transaction history or explicit expectation for future interactions (McKnight et al. 1998). Where the initial level of trust stands can substantially affect the strategies and actions necessary to develop a sustainable trusting relationship. After studying spontaneous trust, we investigate an additional set-up with repeated interactions to demonstrate whether and how the dynamics of trust and trustworthiness may differ between China and the U.S. when there is prospect to develop long-term relationships. Since engaging in long-term interactions is an important prerequisite for the norm of *guanxi* (i.e., close interpersonal connections) in China's business society, examining individuals' behavioral dynamics pertinent to trust in repeated interactions can shed light on how this norm may be induced in supply chains involving Chinese partners.

Some of our contributions to the literature are as follows. First, we study strategic supply chain interactions and information sharing with a cross-country perspective. Neither analytical nor experimental studies on supply chain interactions in the literature have investigated the impact of culture on strategic decisions made by supply chain members. This paper takes the first step to manifest the crucial role of culture in affecting supply chain interactions when information sharing is critical. To this end, we employ an experimental design that allows us to study trust and trustworthiness in strategic information sharing among geographically distant and culturally heterogeneous individuals. Cross-country experiments are rare even in the vast field of behavioral and

experimental economics because conducting such experiments has been technically difficult and costly (Kuwabara et al. 2007). Our experimental design and protocol advance the experimental methodology to enable research on trust, and more broadly, social preferences, in real-time cross-country interactions. Our results offer new empirical insights for how trust and trustworthiness differ between collectivist societies (e.g., China) and individualistic ones (e.g., the U.S.). Second, we characterize and quantify how individuals' countries of origin affect trust, trustworthiness, and the resulting supply chain efficiency from forecast information sharing. We show if, when, and why our Chinese participants demonstrate different trusting and trustworthy behavior than their US counterparts. Our results help to identify conditions under which Chinese individuals exhibit high levels of trust and trustworthiness. For example, we discuss how repeated interactions induce the *guanxi* norm as a guidance of behavior and successfully sustain trust and trustworthiness among Chinese. Finally, our design also allows us to study whether the pull-to-center bias that individuals exhibit when solving a decision problem under uncertainty (i.e., the newsvendor problem[1]) is pervasive in two distinct populations.

## 2. Literature Review

Trust has been studied across different disciplines. One group of researchers investigate how trust is determined by people's attitudes towards risk (Ben-Ner and Putterman 2001), expectation of trustworthiness (Eckel and Wilson 2004, Ashraf et al. 2006), preferences for equality (Hong and Bohnet 2007), and aversion to betrayal (Bohnet and Zeckhauser 2004). Another group focuses on examining the impacts of demographic elements and social status on trust (Croson and Buchan 1999, Gächter et al. 2004, Holm and Nystedt 2005, Hong and Bohnet 2007). These studies primarily use variations of the trust game (Kreps 1990) or the investment game (Berg et al. 1995) to experimentally investigate *spontaneous* trust among strangers in one-time interactions. Other studies adopt a long-term perspective and examine how trust can grow or decline in repeated interactions (Lewicki and Bunker 1995, Doney and Cannon 1997, Engle-Warnick and Slonim 2006). Researchers have advocated developing structurally new games and experimental design to further understand the role of trust in human decision making (Camerer 2003). Following these calls, Özer et al. (2011) introduce the forecast information sharing game to study how trust affects strategic demand forecast sharing in a supply chain.

In parallel, research on trust increasingly gains an international perspective that focuses on understanding how cultural and institutional factors affect trust in different countries or regions

[1] The newsvendor problem represents a basic setting of inventory management in which a newsvendor needs to determine how many newspapers to stock to satisfy uncertain demand at a given day. It is a convex optimization problem that captures the tradeoff between the overage cost of stocking excess inventory and the underage cost of missing demand. The pull-to-center bias describes a systematic behavioral phenomenon observed in newsvendor experiments that individuals often stock closer to the mean demand (the center) than they should optimally.

(La Porta et al. 1997, Doney et al. 1998, Hagen and Choe 1999, Zak and Knack 2001, Wicks and Berman 2004, Bohnet et al. 2010). For example, Knack and Keefer (1997) use data from the World Values Survey[2] to show that trust is generally higher in societies with high income equity, better education, and homogeneous ethnicity. Child and Möllering (2003) show with industry surveys that foreign companies' confidence in Chinese institutions and their active engagement in transferring business practices to the Chinese subsidiaries both positively impact their trust towards the Chinese staff. Bohnet et al. (2008) compare the trust game with a risky dictator game to show that one determinant of trust, betrayal aversion, is a robust phenomenon among different societies. The trust game and the investment game have been replicated in different countries to demonstrate country-level differences in trust and trustworthiness (e.g., Buchan et al. 2002, Willinger et al. 2003, Buchan et al. 2006, Bohnet et al. 2008). In a recent survey, Schoorman et al. (2007) have called for more research to enhance our comprehension of the linkage between culture and trust.

Our contributions to the literature on trust are threefold. First, we employ an experimental design that allows us to study country-level differences in trust and trustworthiness in the context of strategic information communication. To do so, we modify the forecast information sharing game introduced by Özer et al. (2011) in two important aspects: (i) We include a control task to isolate the effects of trust from the effects of individual capability and supply chain factors, and (ii) we implement an online protocol to ensure that participants are convinced of being engaged in real-time interactions with geographically distant individuals. Most prior research on country effects in trust are based on comparing results obtained from each country separately. Thus, they cannot answer how trust may be affected when individuals from different countries interact with each other. Only a handful of studies examine the role of trust in inter-country interactions (Yamagishi et al. 2005, Kuwabara et al. 2007). We advance the experimental methodology to enable rigorous investigation of trust in *real-time cross-country* interactions in addition to within-country interactions. Second, our results add new empirical evidence to the ongoing debate of whether and how trust and trustworthiness differ between collectivist societies and individualistic ones. Third, the current limited studies of trust in inter-country interactions primarily compare Japanese with Westerners. Although Japan and China share similar cultural roots, the institutional and market environments in the two countries are substantially different that trust can manifest distinctively (Takahashi et al. 2008). By focusing on China–U.S. interactions, our results offer concrete managerial implications for supply chain management in one of the most important cross-border exchange relationships.

Forecast information sharing is among the most active and important research areas in operations management. Researchers have studied forecast sharing in both non-strategic and strategic settings.

[2] The World Values Survey has been repeatedly conducted in over 87 societies since 1981. A subset of questions in this survey are used to elicit individuals' general trust attitudes. See http://www.worldvaluessurvey.org/.

Studies with a non-strategic perspective focus on quantifying the values of forecast sharing in a supply chain assuming that supply chain members are willing to fully cooperate with each other (e.g., Lee et al. 2000, Aviv 2003). In contrast, studies with a strategic perspective focus on the impact of incentive conflicts on the effectiveness of forecast sharing. They aim to develop contracts to align the pecuniary incentives of different supply chain members and to identify supply chain conditions under which credible forecast information sharing is ensured (e.g., Cachon and Lariviere 2001, Özer and Wei 2006, Ha and Tong 2008, Ren et al. 2010, Shin and Tunca 2010, Kurtuluş et al. 2012, Gümüş 2013, Kong et al. 2013). Özer et al. (2011) are the first to examine forecast sharing from a behavioral perspective and demonstrate that individuals' inherent propensity to trust and to be trustworthy improves the efficacy of forecast sharing. A recent study by Hyndman et al. (2013) also shows that partially truthful communication about private demand forecasts can help supply chain partners coordinate capacity decisions. We contribute to the forecast sharing literature by introducing a cross-country perspective. We determine and quantify the impact of supply chain members' countries of origin on trust, trustworthiness, and the resulting information sharing efficiency. Our results provide evidence that differences in the supply chain members' countries of origin could be another reason for demand information distortion in a supply chain, in addition to the reasons discussed in Lee et al. (1997). Further, we deepen the experimental analysis of forecast sharing in repeated interactions to obtain insights regarding how the dynamics of trust and trustworthiness in long-term relationships may differ between China and the U.S. This paper also adds to the recent literature that considers country effects in operations and supply chain management (see Tsui et al. 2004, Metters et al. 2010 for reviews).

Behavioral operations management is a fast growing research area. Since Sterman (1989), researchers have studied behavioral issues in various operations contexts, ranging from the bull-whip effect (Steckel et al. 2004, Croson and Donohue 2006), newsvendor decisions (Schweitzer and Cachon 2000, Bolton and Katok 2008, Su 2008), procurement auctions (Cason et al. 2011), supply chain contracting (Cui et al. 2007, Lim and Ho 2007, Chen et al. 2008, Ho and Zhang 2008, Katok and Wu 2009, Kalkanci et al. 2011), to information sharing in a supply chain (Özer et al. 2011). We refer readers to Bendoly et al. (2006, 2010) and Gino and Pisano (2008) for comprehensive reviews. These studies focus on examining whether and how human decisions deviate from the predictions by analytical models, as well as identifying the underlying behavioral causes for the observed deviations. The experiments involved in these studies were mainly conducted in western countries.[3] We contribute to the behavioral operations management literature in two aspects. First,

[3] We encountered only two papers that carry out the newsvendor experiment in China (Feng et al. 2011, Cui et al. 2013). Both papers only conducted experiments in China and compare their results with prior studies involving U.S. participants. Since these studies were conducted independently from the U.S. studies in comparison, their conclusions may be affected by the lack of proper controls on experimenter, language, currency, and subject pool effects (see Roth et al. 1991).

we study behavioral issues in strategic supply chain interactions with a cross-country perspective. We investigate country-level differences between China and the U.S. regarding trusting and trustworthy relationships between supply chain members in information-critical transactions. Our results manifest the crucial role that culture plays in influencing supply chain interactions. Second, by conducting controlled experiments with the exact same design and procedure in both China and the U.S., we document new evidence that the pull-to-center bias observed in prior newsvendor experiments is robust in two distinct countries.

## 3. The Forecast Information Sharing Experiment

We design our experimental set-up based on the forecast information sharing game introduced by Özer et al. (2011). Consider a two-tier supply chain with one supplier (he) and one retailer (she).[4] The supplier produces a product and sells to the retailer, who then sells the product to the end customers. Because of her proximity to the customers, the retailer has better demand forecast information. The supplier wants to solicit this information to plan his production before demand is realized. The final customer demand is random and is modeled as $D = \xi + \epsilon$. The notation $\xi$ represents the retailer's private forecast information. It is deterministically known to the retailer, whereas the supplier only knows that it is a random variable distributed on $[\underline{\xi}, \bar{\xi}]$ with cumulative distribution function (c.d.f.) $F(\cdot)$ and probability density function (p.d.f.) $f(\cdot)$. The notation $\epsilon$ is the market uncertainty of the product. Both parties only know that it is a zero-mean random variable distributed on $[\underline{\epsilon}, \bar{\epsilon}]$ with c.d.f. $G(\cdot)$ and p.d.f. $g(\cdot)$. We assume $\underline{\xi} + \underline{\epsilon} > 0$ to ensure that demand is positive.

The sequence of events is as follows: (i) The retailer observes her private forecast information $\xi$ and submits a report $\hat{\xi}$ to the supplier; (ii) the supplier receives the report $\hat{\xi}$ and produces $Q$ units of the product at a unit cost of $c$; (iii) demand $D$ is realized and the retailer purchases $\min(D, Q)$ from the supplier at a unit wholesale price of $w$; (iv) the retailer sells the product to the end customers at a unit retail price of $r$ and both parties' profits are realized. The game set-up and information structure are common knowledge. Since we are primarily interested in the information communication between the two parties, we regard all the price and cost parameters as exogenous and assume $r \geq w \geq c \geq 0$ to ensure profitable production.

Three key aspects affect the outcome of this forecast sharing game. First, the retailer's trustworthiness impacts the extent of forecast manipulation in her report. A fully trustworthy retailer is willing to credibly share her forecast with the supplier due to a potentially high mental cost of deception. A non-fully trustworthy retailer, however, may distort her forecast for her own interest.

[4] This set-up also applies to other scenarios of a supply chain dyad (e.g., supplier-manufacturer or manufacturer-retailer) in which the downstream member has better demand information than the upstream one and the upstream production lead time is longer than the downstream delivery lead time.

Second, the supplier's trust in the retailer's report affects the way he infers information from the report to make his production decision. A fully trusting supplier regards the report to be truthful and determines the production quantity assuming that the forecast is equal to the report. A non-fully trusting supplier, however, may disregard or use the report when determining the production quantity based on how much he trusts the report. Third, the supplier makes the production decision before the final customer demand is realized. After inferring any information (or not) from the report, the supplier needs to find the production quantity that optimally trades off the cost of building excess inventory (i.e., the overage cost) with the cost of not meeting demand (i.e., the underage cost). This decision problem is known as the newsvendor problem in operations management. All three aspects – trust, trustworthiness, and cognitive ability – collectively determine the outcome of the game and the resulting efficiency of the supply chain.[5] Appendix A characterizes the supplier's and retailer's expected profit functions in detail.

Our primary goal in this paper is to investigate whether and how the supply chain members' *countries of origin* can affect the roles of trust and trustworthiness in strategic forecast information sharing. Note that even when the retailer is fully trustworthy and the supplier is fully trusting, the outcome of the forecast sharing game is still impacted by the supplier's cognitive ability in solving the newsvendor problem. Therefore, we introduce an important task to isolate the effects of trust and trustworthiness on forecast sharing from the effect of participants' cognitive ability. Specifically, we include the "newsvendor task" as a control task in each treatment. In the newsvendor task (Task 1), the retailer's role is computerized and all participants play the role of supplier. The computerized retailer automatically communicates the actual value of her forecast $\xi$ to every participant, and all participants know the truth-telling strategy followed by the computerized retailer. Each participant then determines the production quantity with the credibly shared forecast. Conversely, in the "forecast sharing task" (Task 2), participants take the role of either supplier or retailer. They interact with each other as in the forecast sharing game under random and anonymous pairing. For the suppliers, the only difference between Tasks 1 and 2 is that they make production decisions based on the true $\xi$ in Task 1 versus based on the report $\hat{\xi}$ sent by human retailers in Task 2. The production decision in Task 1 captures the supplier's capability to solve for the optimal production

[5] The standard game theoretic analysis of the forecast sharing game assumes absolutely no trust or trustworthiness between the supply chain members. Özer et al. (2011) show that the only perfect Bayesian equilibrium in this setting is uninformative: The retailer cannot convey any useful information in the report and the supplier disregards the report when determining the production quantity. However, they also experimentally show that a "continuum" of trust and trustworthiness exists between the supply chain members. On one hand, the retailer's report conveys useful information about her forecast, the supplier trusts the report to some extent and determines the production quantity accordingly, leading to informative forecast communication. On the other hand, the retailer still manipulates her forecast to some extent and the supplier does not fully trust the report; i.e., forecast sharing is not fully credible. The existence of this continuum between full trust–trustworthiness and none improves the resulting supply chain efficiency.

quantity when the forecast is credibly shared. The production decision in Task 2 captures the above aspect and in addition the supplier's trust towards the retailer's report. Therefore, the difference in the production decisions between Tasks 1 and 2 elicits the effect of trust.

The above set-up allows us to study country-level differences between China and the U.S. regarding spontaneous trust, trustworthiness, and the resulting information sharing behavior in the supply chain when there is no prospect of long-term interactions. In particular, we study two within-country supply chains (in which both the supplier and the retailer are from the same country) and two cross-country supply chains (in which the supplier and the retailer are from different countries). We seek to answer the following questions: (i) How do trust and trustworthiness of Chinese individuals differ from those of U.S. ones? (ii) Are trust and trustworthiness of Chinese individuals affected by their partners' countries of origin? How about the U.S. individuals? (iii) How much does country of origin impact the resulting supply chain efficiency? We examine these questions under both high and low production costs (corresponding to high and low vulnerability entailed by trusting) to test the robustness of our conclusions.[6]

### 3.1. Measuring Trust and Trustworthiness in Forecast Information Sharing

In our setting, the retailer's trustworthiness is related to how much she distorts her forecast in the report. A more trustworthy retailer tends to distort her forecast to a lesser extent, and a fully trustworthy retailer shares her forecast with the supplier credibly. Given that the underage risk of not meeting demand is salient for the retailer, a non-fully trustworthy retailer has an incentive to inflate her forecast. Hence, we use the *average forecast inflation*, i.e., $(\overline{\hat{\xi}-\xi})$, to measure the retailer's trustworthiness, with a value of zero indicating full trustworthiness and a higher value indicating lower trustworthiness.

Regarding trust, we first consider the supplier's production adjustment from the report in Task 2, i.e., $(Q_2-\hat{\xi})$ where $Q_2$ is the production quantity in Task 2. Note that both the supplier's cognitive ability in solving the newsvendor problem and his trust towards the retailer's report impact this production adjustment. By our design, the cognitive factor is also captured by the supplier's production adjustment in Task 1, i.e., $(Q_1-\xi)$ where $Q_1$ is the production quantity in Task 1. If the supplier fully trusts the report in Task 2, the production adjustments in both tasks, $(Q_2-\hat{\xi})$ and $(Q_1-\xi)$, will be similar on average. In contrast, if the supplier does not fully trust the report in Task 2, he will expect a certain level of forecast inflation by the retailer. As a result, he will first subtract an amount from the report to infer the actual forecast, and then determine $Q_2$ based on similar production adjustment as in Task 1. This thought process is supported by the

[6] Özer et al. (2011) show with U.S. participants that the production cost, which measures the level of risk or vulnerability the supplier endures by trusting the retailer's report, has a significant impact on the levels of trust and trustworthiness in the supply chain.

participants' response to the postexperiment survey. In this case, $(Q_2-\hat{\xi})$ will be lower than $(Q_1-\xi)$ on average. Therefore, we use the *average production adjustment difference*, $(\overline{Q_1-\xi})-(\overline{Q_2-\hat{\xi}})$, to measure the supplier's (dis)trust, with a value of zero indicating full trust and a larger value indicating lower trust.

### 3.2. Hypotheses

We discuss earlier that two aspects impact the outcome of the forecast sharing game: the levels of trust and trustworthiness in the supply chain, and the supplier's capability in solving for the optimal newsvendor decision. We first hypothesize participants' behavior with respect to the second aspect. Prior experiments have consistently shown that when facing uncertain demand, individuals produce too many (few) compared to the optimal quantity if the production cost is high (low) (e.g., Schweitzer and Cachon 2000, Bolton and Katok 2008, Bostian et al. 2008, Lurie and Swaminathan 2009). This observation cannot be explained by risk-averse or risk-seeking preferences, prospect theory, waste aversion, or underestimating opportunity costs. Researchers define this behavior as the pull-to-center bias; i.e., people's decisions are pulled towards the mean demand and away from the optimal decision that balances the cost of holding excess inventory and the cost of missing demand. We expect that this bias exists for both Chinese and U.S. individuals. Thus, we establish the following hypothesis.

HYPOTHESIS 1. *Regardless of country of origin, individuals exhibit the pull-to-center bias in the newsvendor task: Their production quantities fall between the mean demand and the optimal quantity.*

The next set of hypotheses examines how the supply chain members' countries of origin (China versus the U.S.) impact spontaneous trust and trustworthiness in forecast sharing. We first investigate the effect of an individual's own country of origin. Whether individuals from East Asian countries or those from Western countries are more trusting and trustworthy is an ongoing debate that spans multiple disciplines. Social psychologists have long recognized China as a collectivist society whereas the U.S. as an individualistic one (Hofstede 1980, 2001). They posit that average propensity to trust and to be trustworthy is lower in collectivist societies due to a strong in-group bias: Collectivists exhibit high levels of trust and trustworthiness only to their in-groups, i.e., parties that are related to an individual based on kinship or long-term social ties. In contrast, they treat out-group members with suspicion and opportunism (Triandis et al. 1988, Fukuyama 1995, Child 1998, Yamagishi et al. 1998, Chen et al. 2002). These results are mainly obtained with attitudinal surveys. For example, the World Values Survey (2009) shows that over 80% of both Chinese and U.S. individuals report they trust people that they know well, but only 11% of Chinese (versus 40% of U.S. individuals) would trust people that they meet for the first time. Huff and Kelley

(2003) also find evidence that inter-organizational trust is lower in collectivist societies than in individualistic ones due to the in-group bias.

In sharp contrast to the above prediction, however, a large number of experimental studies demonstrate the opposite or no difference. These experimental studies are based on variations of the trust game or the investment game. They show that Chinese participants are either equally or more willing to trust and to reciprocate than U.S. participants (e.g., Buchan et al. 2002, Buchan and Croson 2004, Ho and Weigelt 2005, Buchan et al. 2006). Given these contradicting results, we do not make a priori prediction favoring either direction regarding country-level differences in trust and trustworthiness between China and the U.S. Instead, we propose competing hypotheses and use our observations to empirically test which direction is significant in the context of strategic information sharing. Although the above theories and studies are primarily concerned with trust and trustworthiness within one society, we extrapolate the findings and build the following hypotheses regarding a main effect of an individual's own country of origin.

HYPOTHESIS 2. *(a) Regardless of partner's country of origin, the levels of spontaneous trust and trustworthiness are* lower *in China than in the U.S.; i.e., on average, Chinese retailers induce higher forecast inflation and Chinese suppliers induce larger production adjustment difference than U.S. ones.*

*(b) Regardless of partner's country of origin, the levels of spontaneous trust and trustworthiness are* higher *in China than in the U.S.; i.e., on average, Chinese retailers induce lower forecast inflation and Chinese suppliers induce smaller production adjustment difference than U.S. ones.*

We next discuss how the partner's country of origin may affect an individual's tendency to trust and to be trustworthy in forecast sharing. Recent experiments on intercultural interactions (primarily among different western cultures) show that participants extend more cooperative behavior towards partners of the same nationality than those of different nationalities (Glaeser et al. 2000, Castro 2008). These findings conform to social identity theory (Tajfel and Turner 1979), which posits that intergroup behaviors are largely affected by an individual's perceived membership in a social group. An individual's self-association with a group can generate cooperative intragroup behavior but intensify incentive conflicts in intergroup relations (Akerlof and Kranton 2000). In a cross-country context, a natural and salient social group characteristic is the country of origin. The organization literature indeed shows that different cultural values, expectations, and work practices contribute to conflicts in cross-country work teams and impede collaboration (see Hinds et al. 2011 for a review). Building on these theories, one would expect that in our forecast sharing experiment with one-time interactions, participants are more reluctant to trust or to be trustworthy when interacting with a partner from a different country. We summarize this prediction in the following hypothesis.

HYPOTHESIS 3. *Both U.S. and Chinese individuals show* lower *spontaneous trust and trustworthiness to partners from a different country than to those from the same country. That is, average forecast inflation is higher and average production adjustment difference is larger in cross-country supply chains than in within-country supply chains.*

While the above prediction from social identity theory has been consistently shown among western countries (see Tajfel 2010 for a review), several recent experimental studies that investigate how East Asians behave in intercultural relations with Westerners yield somewhat mixed results. Some find that East Asians cooperate more with East Asians than with Westerners (e.g., Kuwabara et al. 2007, Chuah et al. 2007, Matsumoto and Hwang 2011), whereas others find the opposite (Yamagishi et al. 2005). In Yamagishi et al. (2005)'s Prisoner's Dilemma experiment, Japanese participants cooperate more with Australian partners than with Japanese ones. The authors show that this outcome is consistent with the perception of the Japanese that Australians are more trustworthy. We conjecture that similar stereotyping can play a role in our cross-country supply chains. Since the early 1980s, increasing prevalence of foreign brands and media exposure in China has helped to shape a positive western stereotype among the Chinese population (Batra et al. 2000, Zhou and Hui 2003). Chinese individuals tend to regard Westerners as more open-minded, honest, having higher moral grounds and social status (Bond 1986, Willnat et al. 1997, Huang 2011). These perceptions will likely lead to the Chinese participants being more willing to cooperate with U.S. partners than with Chinese ones. Thus, we make the following hypothesis regarding the Chinese individuals' behavior.

HYPOTHESIS 4. *Chinese individuals show* higher *spontaneous trust and trustworthiness to U.S. partners than to Chinese ones. That is, on average, Chinese retailers induce lower forecast inflation when facing U.S. suppliers, and Chinese suppliers induce smaller production adjustment difference when facing U.S. retailers.*

## 4. Experimental Design and Procedure

Table 1 summarizes the design of our forecast sharing experiment. We used a 2 (self country of origin: China versus the U.S.) × 2 (partner's country of origin: China versus the U.S.) × 2 (high versus low production cost) between-subject design; i.e., we had eight treatments in total. The treatment labels follow the convention $S_iR_jC_k$. The subscript for S (R) indicates the supplier's (retailer's) country of origin, and the subscript for C indicates the cost condition. These treatments represent four supply chain configurations regarding country of origin: a U.S. supply chain, a Chinese supply chain, a Chinese supplier – U.S. retailer supply chain, and a U.S. supplier – Chinese retailer supply chain. We examine two cost conditions to study the robustness of the country effects. Each participant took part in only one of the treatments. In each treatment, participants first made

decisions for 15 rounds in the newsvendor task (Task 1), followed by 15 rounds in the forecast sharing task (Task 2).[7] The participants were informed of the number of rounds in both tasks. For treatments that involve real-time cross-country interactions (i.e., cross-country treatments), participants from one country were assigned the same role and made decisions in that role for all 15 rounds of Task 2. For treatments that involve participants from only one country (i.e., within-country treatments), half of the participants were randomly chosen and assigned the role of retailer at the beginning of Task 2, and the rest were assigned the role of supplier. This role assignment was fixed throughout Task 2.[8] To control for reputation effects, each participant was randomly and anonymously matched with a participant in the other role in each round. Participants were informed that they would *not* play with the same partner in two consecutive rounds.

**Table 1 Experimental Design**

| Condition* | Supplier origin | Retailer origin | Cost | No. of participants | No. of rounds† |
|---|---|---|---|---|---|
| $S_{US}R_{US}C_H$ | U.S. | U.S. | 80 | 20 | 15 |
| $S_{US}R_{CN}C_H$ | U.S. | China | 80 | 24 | 15 |
| $S_{CN}R_{US}C_H$ | China | U.S. | 80 | 24 | 15 |
| $S_{CN}R_{CN}C_H$ | China | China | 80 | 24 | 15 |
| $S_{US}R_{US}C_L$ | U.S. | U.S. | 20 | 24 | 15 |
| $S_{US}R_{CN}C_L$ | U.S. | China | 20 | 24 | 15 |
| $S_{CN}R_{US}C_L$ | China | U.S. | 20 | 22 | 15 |
| $S_{CN}R_{CN}C_L$ | China | China | 20 | 24 | 15 |

In all treatments, $r = 140$, $w = 100$, $\xi$ and $\epsilon$ are uniformly distributed on $[100, 400]$ and $[-75, 75]$. All within-country treatments contained two sessions.
∗: $S_i$ with $i =$ US or CN represents a U.S. or a Chinese supplier; $R_i$ with $i =$ US or CN represents a U.S. or a Chinese retailer; and $C_i$ with $i =$ H or L represents a high or a low cost.
†: Values show the number of rounds in each task.

We conducted all treatments in two computer laboratories, one in the University of Texas at Dallas, TX, U.S., and the other in Tsinghua University, Beijing, China. All within-country treatments were conducted during daytime in the local time zones. All cross-country treatments were conducted at 7:00 p.m. Central Daylight Time in the U.S., corresponding to 8:00 a.m. in China. We used proper controls for multi-country experiments outlined in Roth et al. (1991) to ensure that country effects are not confounded with demographic or background factors (see Appendix B). All participants were undergraduate students with backgrounds in science, engineering, and business. 37% of them were female, they were on average 21 years old (with a standard deviation of 2), and 27% of them were majoring in economics or business related disciplines. The U.S.

[7] One may question the possibility of introducing order effects by having all participants to complete the two tasks in the same order. We choose to use the same order in all treatments to ensure that all supplier participants have the same level of experience in handling the complex newsvendor decision. We determine that order effects do not compromise our conclusions because we mainly focus on cross-treatment comparisons rather than cross-task comparisons, and we use participants' decisions in Task 1 only as a control variable in our regression analysis. In addition, our data do not show evidence that participants from different treatments demonstrate different learning patterns in Task 1.

[8] In the experimental economics literature, both fixed and alternate (randomly or by design) role assignments are commonly used (e.g., Valley et al. 2002, Cai and Wang 2006). We choose to fix participants' roles in our experiment to maintain strong salience of our primary treatment factor, i.e., country of origin.

participants have lived in the U.S. for 17 years on average (with a standard deviation of 7). The Chinese participants have lived in China for 21 years on average (with a standard deviation of 2). Using student participants to study fundamental human behavioral factors (such as trust) is well justified in the experimental economics literature (Friedman and Sunder 1994, Chapter 4). The cross-cultural psychology literature also shows that residence in a country for over 6 years results in an individual's significant adaptation to the cultures and social norms in the country of settlement (Berry 1997, Berry et al. 2006). Therefore, our participants compose a good sample that represents the cultures in the two countries. We subsequently verify that participants' demographic characteristics, academic backgrounds, and work experience are not the driving factors for our conclusions (see §5.3 and Appendix B). The treatments were conducted in English in the U.S. and in Chinese in China. We used back-translation to ensure that the description of the tasks in the two languages is consistent. Two experimenters, both of whom are native Chinese speakers and fluent in English, conducted all treatments with each dedicated to one country. Before conducting the actual treatments, both experimenters ran pilot treatments in the U.S. with no experimenter effect found. This procedure to check for possible experimenter effect is well established and widely used in multi-country experimental studies (e.g., Bohnet et al. 2008). Participants were provided monetary compensation based on the total experimental profits they earned in the experiment plus a show-up fee. We calibrated the payments to achieve compensation parity between the two countries based on the differences in cost of living (UBS 2006) and inputs from local experts. The U.S. participants earned on average 30 U.S. dollars (plus a $10 show-up fee), with a minimum of $15.85 and a maximum of $39.51. The Chinese participants earned on average 100 Chinese yuan (plus a 20 yuan show-up fee), with a minimum of 80 yuan and a maximum of 135 yuan.

All treatments were conducted with the following common protocol. The participants were not allowed to talk among themselves from the time they entered the lab until they left. They made decisions and interacted only through computer terminals during the experiment. The experimental tasks were implemented using the z-Tree software (Fischbacher 2007). Participants worked on first Task 1 and then Task 2. Before each task started, participants read the instructions, answered practice questions, and made decisions for three practice rounds. We pre-generated the random values of $\xi$ and $\epsilon$ for each round of both tasks. Therefore, participants in all treatments faced the same samples of both variables. After participants made their decisions in a round, they were provided feedback. In Task 1, they were shown the value of $\xi$, the realized demand, their production decision, their profits in the current round, and their cumulative profits. In Task 2, they were shown the retailer's report, the supplier's production decision, the realized demand, their profits in the current round, and their cumulative profits. After all rounds were finished, participants were required to complete a postexperiment survey about demographic and academic backgrounds.

Finally, participants collected their payments in private and left the lab. The instructions, sample screenshots, and surveys are available from the authors upon request.

One important aspect of implementing the cross-country treatments is to ensure that participants understand and believe that they would interact with human participants from the other country in Task 2. To do so, we implement the following protocol. After Task 1 was finished, the two experimenters connected via an online video-chat program and showed the video to the participants in their respective lab. The experimenter in China first introduced herself to the U.S. participants. She then showed the lab and the presence of the Chinese participants to the U.S. ones via the camera. She also explained that the Chinese participants were all undergraduate students, had finished the same Task 1, knew that they would interact with U.S. participants in Task 2, and were provided the same instructions for this task. Afterwards, the experimenter in the U.S. repeated the same process for the Chinese participants. Participants in one country observed the presence of the participants in the other country but could not talk to them. Based on the participants' response to the postexperiment survey, they indeed understood and believed that they had interacted with human participants in the other country in Task 2. For example, participants repeatedly stated that playing with students from the other country made them think about the cultural differences between the two populations.

## 5. Experimental Results

Table 2 presents the summary statistics of the participants' decisions and our measures of trust and trustworthiness. We highlight three preliminary observations. First, the production decisions in Task 1 by the U.S. participants are similar to those by the Chinese ones under either cost condition, implying similar cognitive ability between the two pools. Second, the correlations between $\hat{\xi}$ and $\xi$ and between $Q_2$ and $\hat{\xi}$ are strong and significantly positive in all treatments ($t$ tests, $p < 0.01$). This result suggests that the retailers conveyed useful information about their private forecasts to the suppliers, and the suppliers relied on the retailers' reports to make production decisions. Further, both the average forecast inflation ($\overline{\hat{\xi} - \xi}$) and production adjustment difference ($\overline{Q_1 - \xi}$) − ($\overline{Q_2 - \hat{\xi}}$) are significantly positive (Wilcoxon signed rank tests, $p < 0.01$). Recall from §3.1 that if the participants have full trust and trustworthiness, the above two terms achieve values of zero. Thus, these results provide evidence that Özer et al. (2011)'s finding of a continuum between full trust–trustworthiness and none existing in forecast sharing is robust in both countries. Finally, we observe that Chinese retailers inflate forecasts more regardless of the partners' countries of origin and the production cost. The production adjustment difference by Chinese suppliers also tends to be larger than that by U.S. ones. These initial observations indicate that the supply chain members' countries of origin have an impact on trust and trustworthiness in forecast sharing. In the following sections, we formally test the hypotheses established in §3.2.

**Table 2 Summary Statistics: Mean, [Median], (Standard Deviation)**

| | Task 1* | | Task 2† | | | |
|---|---|---|---|---|---|---|
| **Treatment** | $Q_1$ **by U.S.** | $Q_1$ **by China** | $(\overline{\hat{\xi}-\xi})$ | $\mathbf{Cor}(\hat{\xi},\xi)$ | $(\overline{Q_1-\xi})-(\overline{Q_2-\hat{\xi}})$ | $\mathbf{Cor}(Q,\hat{\xi})$ |
| $S_{US}R_{US}C_H$ | 232 [220] (91) | – | 21 [14] (35) | 0.92 | 21 [10] (40) | 0.92 |
| $S_{US}R_{CN}C_H$ | 249 [236] (91) | 234 [221] (95) | 42 [39] (36) | 0.91 | 36 [34] (36) | 0.89 |
| $S_{CN}R_{US}C_H$ | 238 [230] (98) | 235 [225] (87) | 27 [13] (56) | 0.78 | 36 [26] (49) | 0.82 |
| $S_{CN}R_{CN}C_H$ | – | 240 [229] (88) | 48 [38] (47) | 0.83 | 56 [45] (59) | 0.68 |
| $S_{US}R_{US}C_L$ | 279 [271] (88) | – | 31 [16] (44) | 0.86 | 20 [19] (41) | 0.87 |
| $S_{US}R_{CN}C_L$ | 272 [266] (87) | 277 [268] (86) | 61 [53] (47) | 0.84 | 33 [25] (51) | 0.76 |
| $S_{CN}R_{US}C_L$ | 276 [265] (90) | 272 [264] (86) | 12 [10] (41) | 0.88 | 17 [11] (29) | 0.94 |
| $S_{CN}R_{CN}C_L$ | – | 270 [260] (85) | 62 [60] (48) | 0.83 | 78 [73] (50) | 0.79 |

∗: All participants made decisions as suppliers in Task 1, regardless of their roles in Task 2.

†: Label "Cor" means correlation. All correlations are significant with $p < 0.01$.

### 5.1. Do Chinese and U.S. Participants Make Different Newsvendor Decisions?

We first discuss participants' production decisions in the newsvendor task (i.e., Hypothesis 1). We compare their production decisions with the mean demand (i.e., the actual value of $\xi$ in each round) and the optimal production quantity for each treatment. In our experiment, the optimal production quantities are $Q_1^*(\xi) = \xi - 45$ and $\xi + 45$ for a high and a low production cost (see Appendix A). Figure 1 visualizes the comparison for one of the treatments. All other treatments demonstrate similar patterns. For each round, we compute the average production quantities by Chinese and U.S. participants respectively and plot them in comparison to the mean demand and the optimal quantity. We observe that the average production quantities by both populations fall between the mean demand and the optimal quantity. This observation is confirmed by the Wilcoxon signed rank tests. Regardless of country of origin, participants produce more than the optimal quantity but less than the mean demand when the production cost is high, whereas they produce less than the optimal quantity but more than the mean demand when the cost is low (all are significant with $p < 0.05$). Thus, we document new evidence that the pull-to-center bias observed in earlier newsvendor experiments is robust among Chinese and U.S. individuals, supporting Hypothesis 1.

Our experiment also enables us to investigate whether Chinese and U.S. participants perform differently in solving a newsvendor problem, as well as how the pull-to-center bias reacts to changes in the production cost. We observe in Figure 1 that the average production quantities by Chinese and U.S. participants almost overlap with each other, suggesting similar pull-to-center biases across the two countries. To formally answer the above questions, we use random-effects general linear models (GLMs) to compare the following two terms across treatments: participants' production quantities and adjustment scores. The adjustment score is defined by Schweitzer and Cachon (2000) as $(Q_1 - \xi)/(Q_1^*(\xi) - \xi)$. This score measures the intensity of the pull-to-center bias, with a value closer to 0 indicating a stronger bias and a value closer to 1 indicating a closer-to-optimal decision. Using random-effects GLMs to test treatment effects is a well-established methodology

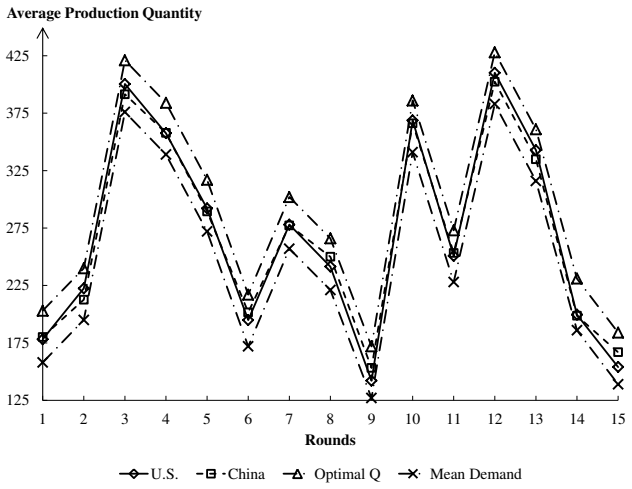



**Figure 1 An Example of Average Production Quantities by Chinese and U.S. Participants in Task 1**

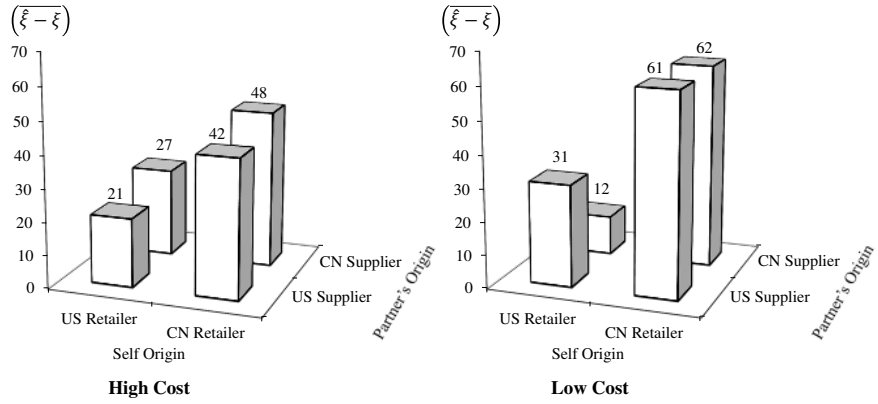
(e.g., Montmarquette et al. 2004, Özer et al. 2011). The detailed regression models and estimates are presented in Appendix D.

We highlight two key findings. First, neither the production quantities nor the adjustment scores differ significantly between Chinese and U.S. participants under either cost. These results are consistent with our graphical observation and suggest that the intensity of the pull-to-center bias is similar in both countries. Second, we observe that in both countries, the participants' adjustment scores are significantly higher with a low production cost ($p < 0.01$). This result indicates that participants make closer-to-optimal newsvendor decisions when the production cost is low, confirming earlier results by Schweitzer and Cachon (2000) and Bolton and Katok (2008).[9] Thus, we conclude that our Chinese and U.S. participants demonstrate similar pull-to-center bias in solving a newsvendor problem, and they make better decisions in a low-cost environment.

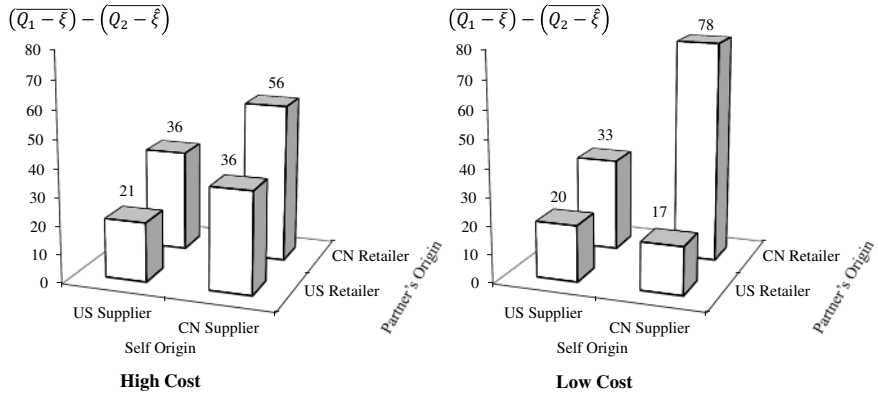
### 5.2. Trust and Trustworthiness in Forecast Information Sharing

Figure 2 visualizes the average forecast inflation (measuring trustworthiness) and production adjustment difference (measuring trust) based on the supply chain members' countries of origin under both cost conditions. We observe differences in behavior with respect to participants' own countries of origin and their partners'. For example, Chinese retailers inflate more than U.S. ones, and Chinese suppliers induce a larger production adjustment difference in most cases. These preliminary observations shed light on potential systematic behavioral differences across the two populations that are pertinent to Hypotheses 2–4. To formally test these hypotheses, we use random-effects GLMs to investigate how the supply chain members' countries of origin affect trust and

[9] Schweitzer and Cachon (2000) show that in their data, the adjustment scores are higher (but not significant) for the low-cost condition than for the high-cost condition. Bolton and Katok (2008) observe that when participants make the newsvendor decisions for 100 periods, they gradually perform better as they gain experience and the improvement is more pronounced under a low cost.




(a) Retailers' Forecast Inflation




(b) Suppliers' Production Adjustment Difference

**Figure 2 Comparing Average Forecast Inflation and Production Adjustment Difference by Country of Origin**

trustworthiness in the forecast sharing task. In particular, we focus on two dependent variables: the retailers' reports $\hat{\xi}$ and the suppliers' production adjustments in Task 2 $(Q_2 - \hat{\xi})$. We analyze our data in the high-cost and low-cost conditions separately due to our focus on country effects. As an alternative approach, we pool all the data and include a dummy variable for the cost condition in the GLMs. All our conclusions hold under the pooled model. Hence, we present the simpler models here. Note that the observed reports are bounded and hence censored by the support of the distribution of $\xi$. Therefore, we use random-effects Tobit models to account for possible censoring. The regression models are as follows:

$$\hat{\xi}^*_{it} = \text{Intercept} + \lambda_{CN} \cdot CN + \lambda_{CNP} \cdot CNP + \lambda_{CNCNP} \cdot CN \cdot CNP + \lambda_{\xi} \cdot \xi_{it} + \lambda_t \cdot t + \delta_i + \varepsilon_{it}, \quad (1)$$

$$(Q_2 - \hat{\xi})_{it} = \text{Intercept} + \lambda_{CN} \cdot CN + \lambda_{CNP} \cdot CNP + \lambda_{CNCNP} \cdot CN \cdot CNP + \lambda_{Q1adj} \cdot (\overline{Q_1 - \xi})_i + \lambda_t \cdot t + \delta_i + \varepsilon_{it}. \quad (2)$$

$\hat{\xi}^*$ is the uncensored latent variable for $\hat{\xi}$ in the Tobit model.[10] The subscripts $i$ and $t$ are participant

[10] We present the regression model and significance results for the latent variable for simplicity. We verify that all our conclusions are valid for the censored $\hat{\xi}$ due to the limited amount of censoring in our data (75 out of 1395 data

and round indices. The two dummy variables *CN* and *CNP* indicate a participant's and his/her partner's country of origin: $CN = 1$ if the participant is from China and 0 otherwise; $CNP = 1$ if the participant's partner is from China and 0 otherwise. We include the variable $CN \cdot CNP$ to investigate possible interaction effects. The variable $t$ is included in both equations to capture possible time trends in participants' decisions. $\xi_{it}$ is included in Equation (1) to examine the dependency between the reports and the forecasts. We include participant $i$'s average production adjustment in Task 1, $(\overline{Q_1 - \xi})_i$, in Equation (2) to control for heterogeneity in the participants' capability to solve the newsvendor problem.[11] Finally, the error terms consist of an individual-specific component $\delta_i$ and an independent component $\varepsilon_{it}$. They enable us to capture individual heterogeneity and the correlation in the decision errors from the same participant (Greene 2012, Chapter 11.5).

The coefficient estimates of Equations (1) and (2) are summarized in Table 3. Note that after controlling for the values of $\xi$ and $(\overline{Q_1 - \xi})$, a lower average value of $\hat{\xi}$ and $(Q_2 - \hat{\xi})$ is equivalent to a lower value of average forecast inflation $(\overline{\hat{\xi} - \xi})$ (i.e., higher trustworthiness) and a larger value of average production adjustment difference $(\overline{Q_1 - \xi}) - (\overline{Q_2 - \hat{\xi}})$ (i.e., lower trust). Thus, the coefficient estimates for the treatment dummies demonstrate how our measures of trust and trustworthiness are affected by the supply chain members' countries of origin. For example, the significantly positive coefficients for *CN* in Equation (1) indicate that Chinese retailers on average inflate forecasts more (i.e., are less trustworthy) than U.S. ones. We next analyze Hypotheses 2–4.

**Table 3 Regression Results for Equations (1) and (2)**

| | Value (standard error) | | | |
|---|---|---|---|---|
| | High cost | | Low cost | |
| **Variable** | $\hat{\xi}^*$ | $Q_2 - \hat{\xi}$ | $\hat{\xi}^*$ | $Q_2 - \hat{\xi}$ |
| Intercept | 58.14 (10.48)$^{**}$ | -0.167 (8.01) | 75.85 (9.27)$^{**}$ | 14.45 (9.85)$^{\dagger}$ |
| *CN* | 20.20 (13.76)$^{\dagger}$ | -17.28 (10.14)$^{\dagger}$ | 30.48 (12.14)$^{**}$ | 1.43 (10.15) |
| *CNP* | 5.06 (13.76) | -10.27 (10.27) | -18.44 (12.41)$^{\dagger}$ | -14.94 (9.92)$^{\dagger}$ |
| $CN \cdot CNP$ | 0.51 (19.32) | -8.54 (14.08) | 19.16 (17.43) | -46.37 (14.23)$^{**}$ |
| $t$ | 1.37 (0.28)$^{**}$ | -3.48 (0.34)$^{**}$ | 1.14 (0.30)$^{**}$ | -3.32 (0.30)$^{**}$ |
| $\xi$ | 0.82 (0.01)$^{**}$ | – | 0.79 (0.02)$^{**}$ | – |
| $(\overline{Q_1 - \xi})$ | – | 0.55 (0.20)$^{**}$ | – | 0.61 (0.23)$^{**}$ |

Notes. "–" means the corresponding variable is not present in the regression.
$**: p < 0.01$; $*: p < 0.05$; $\dagger: p < 0.1$; $p$ values are derived from $t$ tests.

points are censored). See Greene (2012) Chapter 19 (p. 849) for more details on obtaining marginal effects in Tobit models.

[11] Alternatively, one can control this capability by including $(Q_1 - \xi)_i$ when $\xi$ is closest to $\hat{\xi}$ in the current data point. We thank the anonymous reviewer for this suggestion. We estimate this alternative model and obtain identical results.

**5.2.1. How Do Spontaneous Trust and Trustworthiness Compare between China and the U.S.?** We first investigate how the levels of spontaneous trust and trustworthiness differ between the Chinese and the U.S. participants (i.e., Hypothesis 2). We observe from Figure 2(a) that Chinese retailers on average inflate forecasts twice as much as U.S. ones, regardless of the suppliers' countries of origin and the production cost. The top panel of Table 4 summarizes the coefficient estimates relevant to these comparisons based on Equation (1). A positive value indicates higher forecast inflation by Chinese retailers. Note that all values are positive and significant under both costs, corroborating our observation in Figure 2(a). Next we observe from Figure 2(b) that the production adjustment difference is larger for Chinese suppliers than for U.S. ones regardless of the retailers' countries of origin and cost (except in the comparison between $S_{CN}R_{US}C_L$ and $S_{US}R_{US}C_L$ where the values are not significantly different). The coefficient estimates summarized in the bottom panel of Table 4 again confirm this observation. A negative value indicates a larger production adjustment difference by Chinese suppliers. These results show that our Chinese participants exhibit lower levels of spontaneous trust and trustworthiness in forecast sharing compared to their U.S. counterparts.[12] The most drastic difference is demonstrated in that both forecast inflation and production adjustment difference are more than twice as high in the Chinese supply chain as in the U.S. one (see Figure 2). Thus, we support Hypothesis 2(a) and reject Hypothesis 2(b). Our result is therefore consistent with the argument by social psychologists that it is more difficult for spontaneous trust and trustworthiness to occur in a collectivist society such as China than in an individualistic one such as the U.S. This phenomenon may be due to a strong in-group bias. In addition, our observation is also in accordance with former findings that a market-oriented society (such as the U.S.) better promotes cooperation in the general population than one with lower levels of market participation (such as China; see, e.g., Ockenfels and Weimann 1999, Henrich et al. 2001).

**5.2.2. Does the Partner's Country of Origin Impact One's Tendency to Trust and to Be Trustworthy?** We next test Hypotheses 3 and 4 to examine the effect of partner's country of origin on an individual's trust and trustworthiness. We start by discussing the retailer's behavior, i.e., trustworthiness. Note from Figure 2(a) and the top panel of Table 5 that the suppliers' countries of origin have a minimal impact on forecast inflation for both U.S. and Chinese retailers. The only exception is in the comparison between $S_{CN}R_{US}C_L$ and $S_{US}R_{US}C_L$ where U.S. retailers inflate less (marginally significant) when facing Chinese suppliers than when facing U.S. ones. The participants' response to the postexperiment survey indicates that the U.S. retailers showed more

[12] This conclusion is also supported by another study of ours in which we conducted two additional sessions for each of the four within-country treatments.

**Table 4 Effects of Self Country of Origin**

| **Retailers' report $\hat{\xi}$** | | | |
|---|---|---|---|
| | | **Value‡ (standard error)** | |
| **Comparison** | **Coefficient** | **High cost** | **Low cost** |
| $S_{CN}R_{CN} - S_{US}R_{US}$ | $\lambda_{CN} + \lambda_{CNP} + \lambda_{CNCNP}$ | 25.77 (13.76)$^{*}$ | 31.20 (12.14)$^{**}$ |
| $S_{US}R_{CN} - S_{US}R_{US}$ | $\lambda_{CN}$ | 20.20 (13.76)$^{\dagger}$ | 30.48 (12.14)$^{**}$ |
| $S_{CN}R_{CN} - S_{CN}R_{US}$ | $\lambda_{CN} + \lambda_{CNCNP}$ | 20.71 (13.44)$^{\dagger}$ | 49.64 (12.41)$^{**}$ |
| **Suppliers' production adjustment $(Q_2 - \hat{\xi})$** | | | |
| | | **Value‡ (standard error)** | |
| **Comparison** | **Coefficient** | **High cost** | **Low cost** |
| $S_{CN}R_{CN} - S_{US}R_{US}$ | $\lambda_{CN} + \lambda_{CNP} + \lambda_{CNCNP}$ | -36.09 (10.11)$^{**}$ | -59.88 (9.92)$^{**}$ |
| $S_{CN}R_{US} - S_{US}R_{US}$ | $\lambda_{CN}$ | -17.28 (10.14)$^{*}$ | 1.43 (10.15) |
| $S_{CN}R_{CN} - S_{US}R_{CN}$ | $\lambda_{CN} + \lambda_{CNCNP}$ | -25.82 (9.91)$^{**}$ | -44.94 (9.88)$^{**}$ |

‡: Values show changes in the dependent variable between the two treatments.
$**: p < 0.01$; $*: p < 0.05$; $\dagger: p < 0.1$; $p$ values are derived from $t$ tests.

care to the suppliers' profits in $S_{CN}R_{US}C_L$ than in other treatments. This result supports earlier findings that Westerners have the inclination to show fairness by being more cooperative when they interact with individuals from other countries (Singh et al. 1998, Hewstone et al. 2002). However, our data suggest that this behavior can be suppressed when the partner's tendency to cooperate is low. This is evidenced by the fact that the U.S. retailers induce similar forecast inflation for both U.S. and Chinese suppliers when the production cost is high (i.e., when the suppliers have a low tendency to trust). Nevertheless, the majority of our data shows that the retailer's behavior is not affected by the supplier's country of origin. Hence, we do not find support for Hypotheses 3 or 4 regarding trustworthiness.

**Table 5 Effects of Partner's Country of Origin**

| **Retailers' report $\hat{\xi}$** | | | |
|---|---|---|---|
| | | **Value‡ (standard error)** | |
| **Comparison** | **Coefficient** | **High cost** | **Low cost** |
| $S_{CN}R_{US} - S_{US}R_{US}$ | $\lambda_{CNP}$ | 5.06 (13.76) | -18.44 (12.41)$^{\dagger}$ |
| $S_{CN}R_{CN} - S_{US}R_{CN}$ | $\lambda_{CNP} + \lambda_{CNCNP}$ | 5.57 (13.44) | 0.72 (12.14) |
| **Suppliers' production adjustment $(Q_2 - \hat{\xi})$** | | | |
| | | **Value‡ (standard error)** | |
| **Comparison** | **Coefficient** | **High cost** | **Low cost** |
| $S_{US}R_{CN} - S_{US}R_{US}$ | $\lambda_{CNP}$ | -10.27 (10.27) | -14.94 (9.92)$^{\dagger}$ |
| $S_{CN}R_{CN} - S_{CN}R_{US}$ | $\lambda_{CNP} + \lambda_{CNCNP}$ | -18.81 (9.64)$^{*}$ | -61.31 (10.11)$^{**}$ |

‡: Values show changes in the dependent variable between the two treatments.
$**: p < 0.01$; $*: p < 0.05$; $\dagger: p < 0.1$; $p$ values are derived from $t$ tests.

In contrast, Figure 2(b) and the bottom panel of Table 5 demonstrate that the supplier's behavior significantly depends on the retailer's country of origin. Particularly, we observe that under both costs, Chinese suppliers induce a smaller production adjustment difference (i.e., are more trusting)

when facing U.S. retailers than when facing Chinese ones. U.S. suppliers also tend to trust U.S. retailers more, though the differences are marginally significant (under a low cost) or not significant (under a high cost). Thus, our data strongly support Hypothesis 4 with respect to trust and weakly support Hypothesis 3 regarding U.S. individuals' trust.

The higher spontaneous trust Chinese suppliers exhibit towards U.S. retailers may be associated with a positive perception of the U.S. population held by Chinese individuals. To test this assertion, we asked participants to evaluate in the postexperiment survey whether Chinese or U.S. individuals are better described as trusting and/or trustworthy in general. They responded through a 7-point Likert scale, with 1 meaning U.S. individuals are best suited with these attributes, 4 meaning no difference, and 7 meaning Chinese are best suited. The Wilcoxon signed rank tests show that the Chinese participants' scores for the attribute of trustworthiness are significantly lower than the midpoint of 4, and both U.S. and Chinese participants' scores for the attribute of trust are significantly lower than 4 ($p < 0.01$). In addition, Chinese participants' scores on both attributes are significantly lower than those of the U.S. participants (Wilcoxon rank sum test, $p < 0.01$). These results demonstrate that the Chinese participants indeed conceptualize a positive stereotype of the U.S. population being more trusting and trustworthy than the Chinese population. Hence, they are more willing to trust U.S. partners. We also note that these scores do not differ across different treatments, nor do they depend on the role of the participants. Thus, the resulting evaluation is not contaminated by treatment effects or the participants' experience in the experiment.

Another plausible explanation for the observed effects of partner's country of origin is that participants adapt their strategies over time as they learn about their partners' preferences. Since U.S. retailers are on average more trustworthy than Chinese ones (as shown in §5.2.1), it is expected that Chinese suppliers will trust U.S. retailers more. To test how much of the observed effect can be explained by this learning, we estimate a set of random-effects GLMs with additional controls on experience effects. Specifically, we examine how the responses from a participant's past partners may affect the participant's current decisions. The retailers' experience is captured by the paired suppliers' production adjustment ($Q_2 - \hat{\xi}$) in prior rounds. Similarly, the suppliers' experience is captured by the difference between received reports and realized demand ($\hat{\xi} - D$) in prior rounds. The detailed models and estimates are discussed in Appendix E. We observe from these additional models that all the aforementioned significant differences persist after controlling for experience effects. Hence, we conclude that possible learning about the partners' preferences has a marginal explanatory power for the observed effects of partner's country of origin. Instead, these effects are likely driven by more fundamental stereotypical perceptions that Chinese hold for U.S. individuals. Further, we note that the Chinese participants' positive perception of the U.S. people indeed matches with the U.S. participants' behavior in the experiment.

**5.2.3. How Much Does Country of Origin Affect Supply Chain Efficiency?** The above discussions demonstrate that supply chain members' countries of origin significantly impact the levels of spontaneous trust and trustworthiness in a supply chain. Our next step is to quantify these impacts with respect to the resulting supply chain efficiency. Supply chain efficiency, $E$, is defined as the total expected supply chain profit given the observed production decision under asymmetric information in proportion to the optimal expected profit of a centralized supply chain.[13] We use the following random-effects GLM to investigate country effects in supply chain efficiency:

$$E_{it} = \text{Intercept} + \lambda_{CN} \cdot CN + \lambda_{CNP} \cdot CNP + \lambda_{CNCNP} \cdot CN \cdot CNP + \lambda_{\xi} \cdot \xi_{it} + \lambda_t \cdot t + \delta_i + \varepsilon_{it}. \quad (3)$$

The subscript $i$ is the index for a pair of participants that form a supply chain. *CN* indicates the supplier's country of origin and *CNP* indicates the retailer's country of origin in the supply chain. Table 6 summarizes the coefficient estimates for Equation (3) and for all pairwise comparisons between any two supply chain configurations. We first observe that a cross-country supply chain is significantly more efficient than a Chinese supply chain under both production costs (Rows 2 and 3 in Table 6). The higher efficiency in $\mathrm{S_{US}R_{CN}}$ than in $\mathrm{S_{CN}R_{CN}}$ is due to higher trust exhibited by U.S. suppliers. The higher efficiency in $\mathrm{S_{CN}R_{US}}$ than in $\mathrm{S_{CN}R_{CN}}$ is jointly driven by the U.S. retailers' higher trustworthiness and the Chinese suppliers' higher trust towards U.S. partners. Conversely, a U.S. supply chain is more efficient than a cross-country supply chain when the production cost is high, but this difference vanishes when the production cost is low (Rows 4 and 5 in Table 6). Note from Tables 4 and 5 that the increase in trust and/or trustworthiness (if any) in a U.S. supply chain compared to a cross-country supply chain is moderate. When the production cost is high, the supplier can be motivated to build much higher inventory given moderately higher trust and trustworthiness, thereby improving the supply chain efficiency. In contrast, when the production cost is low, such an increase in inventory is limited because the supplier naturally produces a large quantity. Hence, the U.S. participants' higher trust and trustworthiness significantly enhance supply chain efficiency compared to a cross-country supply chain only when the production cost is high. Finally, when the levels of trust and trustworthiness between two supply chains differ the most (i.e., $\mathrm{S_{CN}R_{CN}}$ versus $\mathrm{S_{US}R_{US}}$), the low-trust supply chain can suffer from over 10% efficiency loss and hence profit reduction. These results thus highlight and quantify the impact of country of origin on supply chain efficiency through its effect on trust and trustworthiness.

[13] Formally, the supply chain efficiency $E(Q) = [\Pi^s(\xi, Q) + \Pi^r(\xi, Q)]/\Pi^c(\xi, Q^c(\xi))$. Here, $\Pi^s(\xi, Q)$ and $\Pi^r(\xi, Q)$ are the supplier's and retailer's expected profit with private forecast $\xi$ and production quantity $Q$, and $\Pi^c(\xi, Q^c(\xi))$ is the optimal expected profit of the centralized supply chain, as defined in Appendix A.

**Table 6 Regression and Comparison Results on Supply Chain Efficiency**

| | Value (standard error) | | | Value‡ (standard error) | |
|---|---|---|---|---|---|
| **Variable** | **High cost** | **Low cost** | **Comparison & coefficient** | **High cost** | **Low cost** |
| Intercept | 74.91 (3.06)** | 98.94 (1.10)** | $S_{CN}R_{CN} - S_{US}R_{US}$: $\lambda_{CN} + \lambda_{CNP} + \lambda_{CN\cdot CNP}$ | -10.45 (2.58)** | -3.51 (1.09)** |
| *CN* | -4.67 (2.58)* | -0.20 (1.11) | $S_{CN}R_{CN} - S_{CN}R_{US}$: $\lambda_{CNP} + \lambda_{CN\cdot CNP}$ | -5.78 (2.46)** | -3.31 (1.11)** |
| *CNP* | -4.04 (2.58)† | -0.36 (1.09) | $S_{CN}R_{CN} - S_{US}R_{CN}$: $\lambda_{CN} + \lambda_{CN\cdot CNP}$ | -6.41 (2.46)** | -3.15 (1.09)** |
| $CN \cdot CNP$ | -1.74 (3.58) | -2.95 (1.57)* | $S_{CN}R_{US} - S_{US}R_{US}$: $\lambda_{CN}$ | -4.67 (2.58)* | -0.20 (1.11) |
| $t$ | 0.19 (0.21) | -0.05 (0.07) | $S_{US}R_{CN} - S_{US}R_{US}$: $\lambda_{CNP}$ | -4.04 (2.58)† | -0.36 (1.09) |
| $\xi$ | 0.06 (0.01)** | -0.01 (0.00)** | | | |

‡: Values show changes in the dependent variable between the two treatments.
$**: p < 0.01$; $*: p < 0.05$; $\dagger: p < 0.1$; $p$ values are derived from $t$ tests.

### 5.3. Validation of the Experimental Results

This section provides validity checks regarding our experimental results. We discuss earlier that all our conclusions continue to hold when participants' experience is controlled for. In addition, the coefficients for the round index $t$ in Equations (1) and (2) suggest that forecast inflation tends to be higher and the production adjustment difference larger towards the end of the task (see Table 3). Nevertheless, the observed country effects remain valid when these time effects are controlled for. We will revisit the participants' behavioral dynamics when we compare the single-interaction treatments with the repeated-interaction ones in the next section.

We also examine whether the participants' gender, age, majoring in an economics or business related discipline, exposure to game theory, and years of work experience in a business domain affect their forecast sharing behavior. We include these terms as additional independent variables in Equations (1) and (2) to investigate their impacts on trust and trustworthiness. The estimation results (see Appendix E) show that age and majoring in economics or business related disciplines have no effect on the participants' behavior in any treatment. Work experience may induce more trustworthy but less trusting behavior. Under a low cost, male retailers are less trustworthy, and suppliers who know game theory are less trusting. Nevertheless, our observations discussed in §5.2 remain valid after controlling for these demographic and background factors. Finally, we also test Hypotheses 2–4 only with data between rounds 3 and 13 in the forecast sharing task. The reason to do this is to eliminate the potential impacts of initial trials and participants' fatigue towards the end of a treatment on their behavior. We estimate Equations (1) and (2) based on this restricted set of data and observe that all our earlier results continue to hold.

## 6. Repeated Interactions

So far, we have demonstrated that supply chain members' countries of origin have significant impacts on spontaneous trust and trustworthiness within a supply chain. In particular, trust and trustworthiness are more difficult to occur for the Chinese participants than for U.S. ones when there is no prospect for long-term interactions, although interaction with a U.S. partner can induce higher tendency to trust by Chinese individuals. An important question thus arises as to how these

results may vary (or not) when supply chain partners interact repeatedly. Researchers have long recognized the important role of social relations in generating trust. In his seminal paper, Granovetter (1985) establishes the "social embeddedness" argument and postulates that "continuing economic relations often become overlaid with social content that carries strong expectations of trust and abstention from opportunism" (p. 490). Recent studies also demonstrate that *guanxi* contributes to producing trust within and across Chinese business organizations (Farh et al. 1998, Cai et al. 2010). Thus, we conjecture that when supply chain partners interact with each other repeatedly, trust and trustworthiness will substantially increase and the country differences we observe earlier will diminish. In this section, we design and conduct a new set of experimental treatments to examine the following hypotheses.

HYPOTHESIS 5. *(a) Forecast inflation is lower and production adjustment difference is smaller in repeated interactions than in one-time interactions, regardless of the supply chain members' countries of origin.*

*(b) In repeated interactions, forecast inflation and production adjustment difference are similar among Chinese and U.S. individuals, regardless of the partners' countries of origin.*

Özer et al. (2011) also conducted a repeated-interaction version of the forecast sharing game. They focus on the role of information feedback and observe that repeated interactions foster trust among their U.S. participants even if the suppliers cannot observe the actual private forecast information after each interaction. Different than these authors, we continue to focus on country effects in our repeated-interaction treatments. In addition to testing Hypothesis 5, we also investigate whether and how the dynamics of trust and trustworthiness under repeated interactions may differ between U.S. and Chinese individuals.

We conducted four additional treatments as summarized in Table 7. We continue to investigate the four supply chains, $S_{US}R_{US}$, $S_{US}R_{CN}$, $S_{CN}R_{US}$, and $S_{CN}R_{CN}$, with a between-subject design. In the repeated-interaction treatments, we only focus on a high cost. Prior research has shown that trust and trustworthiness are harder to arise when the vulnerability entailed by trusting (measured by the production cost in our context) is high (e.g., Malhotra 2004, Özer et al. 2011), thus leaving more room for improvement with repeated interactions. The treatments were conducted in the same laboratories as before. The same cross-country controls and experimental protocols were implemented. The only difference here is that participants were anonymously and randomly matched into pairs at the beginning of Task 2, and everyone interacted with the same partner in all 15 rounds of the task. Participants were informed of the repeated-interaction set-up and the number of rounds in both tasks. The U.S. participants in these treatments earned \$32.62 (plus a \$10 show-up fee) on average, with a minimum of \$21.62 and a maximum of \$39.13. The Chinese participants earned 107.19 yuan (plus a 20 yuan show-up fee) on average, with a minimum of 80 yuan and a maximum of 140 yuan.

**Table 7 Experimental Design: Repeated Interactions**

| Condition* | Supplier origin | Retailer origin | No. of participants | No. of rounds† |
|---|---|---|---|---|
| $S_{US}R_{US}$REP | U.S. | U.S. | 24 | 15 |
| $S_{US}R_{CN}$REP | U.S. | China | 24 | 15 |
| $S_{CN}R_{US}$REP | China | U.S. | 24 | 15 |
| $S_{CN}R_{CN}$REP | China | China | 24 | 15 |

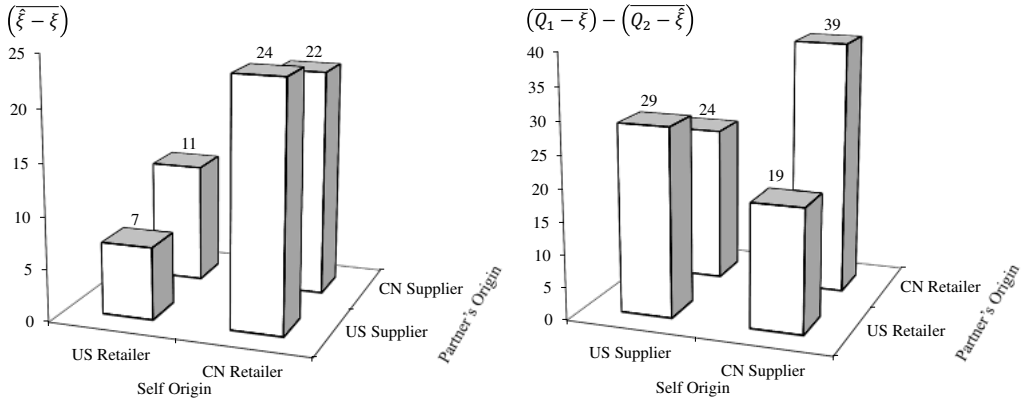
In all treatments, $r = 140$, $w = 100$, $c = 80$, $\xi$ and $\epsilon$ are uniformly distributed on $[100, 400]$ and $[-75, 75]$.
∗: "REP" represents repeated interactions. †: The number shows the number of rounds in each task.

## 6.1. Summary Effects of Repeated Interactions

Table 8 presents the summary statistics and Figure 3 visualizes the average levels of trust and trustworthiness in each of the repeated-interaction treatments. We first note that the participants' production decisions in Task 1 are similar across the two countries and also similar to those in the earlier treatments (see Table 2). These results further confirm that the two populations show similar pull-to-center bias in their newsvendor decisions. Comparing Figures 2 and 3, we note that average forecast inflation is lower and production adjustment difference is smaller under repeated interactions than under one-time interactions.

**Table 8 Repeated-Interaction Treatment Summary Statistics: Mean, [Median], (Standard Deviation)**

| | Task 1 | | Task 2 | |
|---|---|---|---|---|
| Treatment | $Q_1$ by U.S. | $Q_1$ by China | $(\overline{\hat{\xi} - \xi})$ | $(\overline{Q_1 - \xi}) - (\overline{Q_2 - \hat{\xi}})$ |
| $S_{US}R_{US}$REP | 240 [228] (91) | – | 7 [0] (53) | 29 [15] (51) |
| $S_{US}R_{CN}$REP | 241 [230] (90) | 243 [228] (88) | 24 [16] (29) | 24 [7] (52) |
| $S_{CN}R_{US}$REP | 244 [235] (86) | 242 [230] (91) | 11 [3] (28) | 19 [9] (36) |
| $S_{CN}R_{CN}$REP | – | 239 [222] (92) | 22 [13] (27) | 39 [31] (44) |




(a) Retailers' Forecast Inflation

(b) Suppliers' Production Adjustment Difference

**Figure 3 Average Forecast Inflation and Production Adjustment Difference in Repeated-Interaction Treatments**

To formally test Hypothesis 5, we follow the same approach as in §5.2 and compare the data from these new treatments with the data from the single-interaction treatments under a high cost. The

detailed regression models and parameter estimates are discussed in Appendix F. The comparisons relevant to Hypothesis 5 are summarized in Table 9. We highlight three observations. First, except for the U.S. supply chain,[14] repeated interactions indeed lead to higher trust, trustworthiness, and supply chain efficiency in the other three supply chains (the top panel of Table 9). Hence, Hypothesis 5(a) is supported. Second, consistent with Hypothesis 5(b), the supply chain members' countries of origin have a diminished effect on both forecast inflation and production adjustment difference under repeated interactions (the bottom panel of Table 9). In other words, the presence of repeated interactions enhances trust and trustworthiness among Chinese participants to a level similar to those among U.S. participants. Nevertheless, Chinese suppliers continue to trust U.S. partners more than they trust Chinese ones, as shown in the significantly negative difference in $(Q_2-\hat{\xi})$ between $\mathrm{S_{CN}R_{CN}REP}$ and $\mathrm{S_{CN}R_{US}REP}$. To evaluate the impact of repeated interactions on supply chain efficiency, we compare the efficiency *gain* across different supply chain configurations. An efficiency gain is defined as the supply chain efficiency under repeated interactions minus that under one-time interactions. We observe that a supply chain involving Chinese partner(s) gains substantially more efficiency under repeated interactions than a U.S. supply chain (Rows 1–3, last column in the bottom panel of Table 9). For example, the average efficiency gain in the Chinese supply chain from one-time to repeated interactions is over 14% more than that in the U.S. supply chain. This observation suggests that the Chinese participants benefit more from a long-term relationship than U.S. ones do.

**Table 9 Effects of Repeated Interactions**

| **Repeated interactions versus one-time interactions** | | | |
|---|---|---|---|
| | **Value‡ (standard error)** | | |
| **Comparison** | $\hat{\boldsymbol{\xi}}$ | $\boldsymbol{Q_2-\hat{\xi}}$ | **Efficiency (%)** |
| $\mathrm{S_{US}R_{US}REP-S_{US}R_{US}C_H}$ | -11.63 (11.40) | -8.18 (11.10) | -3.30 (2.98) |
| $\mathrm{S_{CN}R_{CN}REP-S_{CN}R_{CN}C_H}$ | -22.84 (11.13)$^{*}$ | 17.38 (10.59)$^{\dagger}$ | 11.59 (2.84)$^{**}$ |
| $\mathrm{S_{CN}R_{US}REP-S_{CN}R_{US}C_H}$ | -15.87 (11.13)$^{\dagger}$ | 18.69 (10.64)$^{*}$ | 9.96 (2.83)$^{**}$ |
| $\mathrm{S_{US}R_{CN}REP-S_{US}R_{CN}C_H}$ | -18.92 (11.13)$^{*}$ | 9.66 (10.64) | 5.58 (2.83)$^{*}$ |
| **Effects of country of origin** | | | |
| | **Value‡ (standard error)** | | |
| **Comparison** | $\hat{\boldsymbol{\xi}}$ | $\boldsymbol{Q_2-\hat{\xi}}$ | **Efficiency Gain (%)♯** |
| $\mathrm{S_{CN}R_{CN}REP-S_{US}R_{US}REP}$ | 14.85 (11.13) | -10.04 (10.58) | 14.90 (4.11)$^{**}$ |
| $\mathrm{S_{CN}R_{US}REP-S_{US}R_{US}REP}$ | 1.12 (11.13) | 10.61 (10.58) | 13.26 (4.11)$^{**}$ |
| $\mathrm{S_{US}R_{CN}REP-S_{US}R_{US}REP}$ | 13.21 (11.13) | 5.59 (10.58) | 8.88 (4.11)$^{*}$ |
| $\mathrm{S_{CN}R_{CN}REP-S_{CN}R_{US}REP}$ | 13.73 (11.13) | -20.65 (10.58)$^{*}$ | 1.28 (4.01) |
| $\mathrm{S_{CN}R_{CN}REP-S_{US}R_{CN}REP}$ | 1.64 (11.13) | -15.63 (10.58)$^{\dagger}$ | 6.02 (4.01)$^{\dagger}$ |

‡: Values show changes in the dependent variable between the two treatments.
♯: This column compares the efficiency gain from one-time to repeated interactions between the two supply chains.
$**: p<0.01$; $*: p<0.05$; $\dagger: p<0.1$; $p$ values are derived from $t$ tests.

[14] See Appendix G for a discussion about the impact of individual heterogeneity on the efficiency of the U.S. supply chain.

### 6.2. Dynamics of Trust and Trustworthiness in Repeated Interactions

Here we study whether and how our participants show different time trends and reactions to experience in repeated interactions versus in one-time interactions. We analyze this comparison for each of the four supply chain configurations respectively. To do so, we estimate additional regression models that incorporate interaction terms between the treatment dummies and the round and experience variables. We use the same experience variables as discussed in §5.2.2. The details are presented in Appendix F. Table 10 summarizes the coefficient estimates pertinent to time and experience effects under each supply chain configuration for both one-time and repeated interactions. A positive (negative) coefficient for the round variable implies that the corresponding dependent variable increases (decreases) over time. Similarly, a positive (negative) coefficient for the experience variable implies that the corresponding dependent variable increases (decreases) as the value of the experience variable increases.

**Table 10 Time and Experience Effects in Repeated versus One-Time Interactions**

| | | Value (standard error) | | | |
|---|---|---|---|---|---|
| | | $\hat{\xi}$ | | $Q_2 - \hat{\xi}$ | |
| | **Treatment** | **Round ($t$)** | **Experience ($Q_2 - \hat{\xi}$)** | **Round ($t$)** | **Experience ($\hat{\xi} - D$)** |
| **Within-Country Supply Chains** | $S_{US}R_{US}C_H$ | 1.11 (0.68) | -0.02 (0.07) | -1.79 (0.80)$^{*}$ | -0.07 (0.06) |
| | $S_{US}R_{US}REP$ | 0.82 (0.57) | 0.16 (0.15) | -1.61 (0.71)$^{*}$ | -0.07 (0.04)$^{\dagger}$ |
| | $S_{CN}R_{CN}C_H$ | 1.81 (0.58)$^{**}$ | -0.09 (0.04)$^{*}$ | -3.99 (0.73)$^{**}$ | -0.21 (0.05)$^{**}$ |
| | $S_{CN}R_{CN}REP$ | 0.85 (0.56) | -0.01 (0.06) | -0.68 (0.71) | -0.21 (0.06)$^{**}$ |
| **Cross-Country Supply Chains** | $S_{CN}R_{US}C_H$ | -0.67 (0.57) | -0.12 (0.05)$^{*}$ | -3.56 (0.74)$^{**}$ | -0.15 (0.06)$^{**}$ |
| | $S_{CN}R_{US}REP$ | 1.08 (0.58)$^{\dagger}$ | -0.28 (0.10)$^{**}$ | -2.10 (0.73)$^{**}$ | -0.30 (0.06)$^{**}$ |
| | $S_{US}R_{CN}C_H$ | 1.00 (0.64) | -0.11 (0.07) | -1.69 (0.71)$^{*}$ | -0.23 (0.04)$^{**}$ |
| | $S_{US}R_{CN}REP$ | 0.75 (0.57) | -0.09 (0.07) | -1.45 (0.74)$^{*}$ | -0.31 (0.07)$^{**}$ |

$**: p < 0.01$; $*: p < 0.05$; $\dagger: p < 0.1$; $p$ values are derived from $t$ tests.

We first focus on the within-country supply chains (Rows 1–4 in Table 10). For the U.S. supply chains, we observe that the U.S. retailers do not change their strategies over time in either one-time or repeated interactions, nor do they react to past experiences. The U.S. suppliers, however, decrease their trust over time in both interaction settings (the coefficients for the round variable are significantly negative). In addition, the non-significant or marginally significant coefficients for the experience variable show that the decline in trust is not driven by negative past experience, but rather, a change in the participants' internal preferences as time progresses.

Conversely, the Chinese supply chains exhibit significantly different behavioral dynamics in one-time versus repeated interactions. In one-time interactions, Chinese retailers and suppliers react strongly to past experience, and both trust and trustworthiness decline significantly over time. However, in repeated interactions, the Chinese retailers employ a stationary strategy throughout the course of interactions, and the Chinese suppliers no longer decrease their trust over time.

Nevertheless, the Chinese suppliers continue to punish less trustworthy retailers by lowering trust (as seen in the significantly negative coefficient for the experience variable).

Finally, we observe that for the cross-country supply chains, our participants show similar time and experience effects between one-time and repeated interactions (Rows 5–8 in Table 10). That is, the presence of long-term relationships does not significantly influence the participants' behavioral dynamics in cross-country supply chains. We postulate that this is because when individuals interact with those from a different country, they exert most of their efforts to learn about their partners' preferences and adapt their behavior accordingly. Therefore, the nature of interactions (one-time versus repeated) has a diminished effect on their behavior.

To conclude, our results in the repeated-interaction treatments conform to social psychology theories that a collectivist society such as China is more relation-driven in social interactions than an individualistic society such as the U.S. (Markus and Kitayama 1991). Our observations in behavioral dynamics demonstrate a sharp contrast between Chinese and U.S. participants regarding the role of long-term relationships in affecting trust and trustworthiness. For U.S. participants, whether or not to trust and to be trustworthy are more dependent on intrinsic propensities and not influenced by the presence (or non-presence) of long-term relationships. In contrast, Chinese participants substantially adjust their behavior when there is prospect to build long-term relationships. A possible reason for these behavioral changes may be attributed to the important norm of *guanxi* in the Chinese society. Since engaging in long-term interactions is a prerequisite for building *guanxi*, the presence of repeated interactions likely induces the *guanxi* norm as a guidance of behavior for the Chinese participants. This norm substantially deters opportunistic behavior and motivates trust (Xin and Pearce 1996). Consequently, the Chinese supply chains significantly benefit from long-term relationships as high levels of trust and trustworthiness are successfully sustained among the Chinese participants. These results also echo the Quantum executives' comments (see §1) that ensuring long-term contact with the same manager is essential for western companies to build trust with Asian partners.

## 7. Discussion and Conclusion

This paper investigates the country-level differences in trust and trustworthiness between China and the U.S. in the context of forecast information sharing. In particular, we focus on a two-tier supply chain in which the upstream supplier solicits demand forecast information from the downstream retailer to make production decisions, and the retailer has an incentive to inflate her private forecast. Two key aspects affect the outcome of the strategic interaction between the supplier and the retailer: the levels of trust and trustworthiness in the supply chain and the supplier's capability in finding the optimal production quantity in face of uncertain demand. We employ an experimental design

to disentangle these two aspects and to study (i) whether and how Chinese and U.S. individuals behave differently when solving a complex decision problem under uncertainty (i.e., the newsvendor problem), (ii) the country-level variations between China and the U.S. in trust and trustworthiness and how these variations impact strategic information sharing, (iii) whether and how Chinese and U.S. individuals' tendency to trust and to be trustworthy depends on their supply chain partners' countries of origin, and (iv) how much country of origin impacts the resulting supply chain efficiency. We study both the occurrence of spontaneous trust and trustworthiness in the absence of repeated interactions and the dynamics of trust and trustworthiness in a long-term relationship. This paper is the first to investigate and quantify how supply chain members' countries of origin affect trust, trustworthiness, and the resulting efficacy of strategic information sharing in global supply chains.

We first demonstrate that both Chinese and U.S. participants exhibit similar pull-to-center bias in their newsvendor decisions. In addition, individuals from both countries make closer-to-optimal decisions when the overage risk is lower. We next show that in both countries, participants exhibit a continuum between full trust–trustworthiness and none, leading to informative forecast communication. We observe that when there is no prospect for long-term interactions, our Chinese participants consistently exhibit lower levels of spontaneous trust and trustworthiness than their U.S. counterparts, regardless of the partner's country of origin and the production cost. We quantify that Chinese retailers on average inflate forecast information twice as much as U.S. ones do. Chinese suppliers also rely less on the forecast report when determining the production quantity. Such lower levels of trust and trustworthiness can result in a 10% loss in supply chain profit and efficiency. We posit that the distinct national culture, the institutional environment, and recent societal changes in China all contribute to the difficulty for spontaneous trust to arise in the country. First, the collectivism orientation of the Chinese culture restricts trust and trustworthiness within one's tight social network (i.e., the in-group) which is typically formed based on family ties or long-term friendship. As a result, opportunistic behavior favoring one's own interest often arises in transactions with out-group members (which is the case for most business transactions). Second, the lack of a facilitative government and reliable legal systems that support independent organizations and market participation also impedes the development of trust and trustworthiness among business partners (Rao et al. 2005). Finally, the drastic economic and social reforms occurring in China in the past few decades have caused a momentous emphasis on competition and personal achievements (Egri and Ralston 2004), which also adversely affects overall trust and trustworthiness in the society. Even the government has realized this adverse effect and calls for the construction of a harmonious society (Fan 2006). Our findings highlight the importance for firms to devote more time and effort to maintaining a trusting and cooperating supply chain relationship with their Chinese partners than they do with U.S. ones. In turn, Chinese companies and government should

also proactively cultivate a cooperative mindset among the young generation and establish an environment conducive for efficient inter-organization transactions.

Our comparison between within-country and cross-country supply chains demonstrates that both Chinese and U.S. individuals trust U.S. partners more than Chinese ones, whereas neither population's trustworthiness is affected by their partners' countries of origin. Engaging in repeated interactions significantly enhances trust and trustworthiness among Chinese individuals to similar levels among U.S. ones. Hence, a Chinese supply chain benefits more from long-term relationships by enjoying a higher efficiency gain than a U.S. supply chain. These observations suggest two effective forces that can help to improve trust, trustworthiness, and efficiency in a cross-country supply chain involving Chinese partners in different settings. When the prospect of long-term interactions is less prominent (e.g., when initiating new business relationships with Chinese companies), western companies can rely on and reinforce the Chinese population's positive perception of Westerners as more trusting and trustworthy in general to encourage higher trust by their Chinese partners. Possible actions include offering better-quality products in the Chinese market or establishing better working conditions for Chinese workers in offshore factories. In addition, when long-term engagement is promising (e.g., when initial transactions with a Chinese partner signal the profitability of a longer-term partnership), western companies should secure a long-term relationship (e.g., by signing long-term contracts) to induce the *guanxi* norm and motivate trust from the Chinese partner.

Studying country effects on supply chain interactions and operational performance opens a fruitful avenue for future research. This paper focuses on determining the existence of country differences in certain behavioral phenomena. We hope to stimulate future studies on trust and cooperation in cross-country supply chains. For example, extending the subject pool to industry practitioners can help to verify the robustness of our conclusions. Earlier experiments show that whether managers and students behave differently depends on the context. Most studies with operations management contexts have shown that the two groups behave similarly (see Croson and Donohue 2006, Bolton et al. 2012, Lim and Ho 2007 for examples of beer game, newsvendor, and contracting experiments). In other contexts, experienced professionals are also shown to exhibit similar behavioral biases as observed in lab experiments with student participants (e.g., Cooper et al. 1999, Massey and Thaler 2013), although evidence to the contrary also exist (e.g., Fehr and List 2004). This question remains unexplored in the context of strategic information sharing. In addition, one can extend the repeated-interaction setting to further study the role of *guanxi* in the Chinese business environment. For example, factors such as family/regional ties and gift exchange are important components for the formation and preservation of *guanxi*. Researchers can build on our experimental design and results to examine the dynamics of trust and trustworthiness in a supply chain

when these factors are present versus absent. Relatedly, it is also interesting to study whether Chinese show different levels of trust and trustworthiness towards ethnic Chinese foreign partners (i.e., Chinese immigrants in western countries) versus non-ethnic Chinese ones (i.e., Westerners), and what is the role of *guanxi* in this comparison. To conclude, we believe that related research opportunities are boundless.

## References


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

Arrow, K. 1972. Gifts and exchanges. *Philosophy & Public Affairs* **1**(4) 343–362.

Ashraf, N., I. Bohnet, N. Piankov. 2006. Decomposing trust and trustworthiness. *Experimental Economics* **9**(3) 193–208.

Aviv, Y. 2003. A time series framework for supply chain inventory management. *Operations Research* **51**(2) 210–227.

Barboza, D. 2010. China passes Japan as second-largest economy. *New York Times*, p. B1, August 16.

Barboza, D., B. Stone. 2010. China, where U.S. internet companies often fail. *New York Times*, p. B1, January 16.

Batra, R., V. Ramaswamy, D. Alden, J. Steenkamp, S. Ramachander. 2000. Effects of brand local and nonlocal origin on consumer attitudes in developing countries. *Boston University Law Review* **9**(2) 83–95.

Ben-Ner, A., L. Putterman. 2001. Trusting and trustworthiness. *Boston University Law Review* **81**(3) 523–551.

Bendoly, E., R. Croson, P. Goncalves, K. Schultz. 2010. Bodies of knowledge for research in behavioral operations. *Production & Operations Management* **19**(4) 434–452.

Bendoly, E., K. Donohue, K. Schultz. 2006. Behavior in operations management: Assessing recent findings and revisiting old assumptions. *Journal of Operations Management* **24**(6) 737–752.

Berg, J., J. Dickhaut, K. McCabe. 1995. Trust, reciprocity, and social history. *Games & Economic Behavior* **10**(1) 122–142.

Berry, J. 1997. Immigration, acculturation, and adaptation. *Applied Psychology: An International Review* **46**(1) 5–68.

Berry, J., J. Phinney, D. Sam, P. Vedder. 2006. Immigrant youth: Acculturation, identity, and adaptation. *Applied Psychology: An International Review* **55**(3) 303–332.

Bohnet, I., F. Greig, B. Herrmann, R. Zeckhauser. 2008. Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *American Economic Review* **98**(1) 294–310.

Bohnet, I., B. Herrmann, R. Zeckhauser. 2010. Trust and the reference points for trustworthiness in Gulf and Western countries. *Quarterly Journal of Economics* **125**(2) 811–828.

Bohnet, I., R. Zeckhauser. 2004. Trust, risk and betrayal. *Journal of Economic Behavior & Organization* **55**(4) 467–484.

Bolton, G., E. Katok. 2008. Learning-by-doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manufacturing & Service Operations Management* **10**(3) 519–538.

Bolton, G., A. Ockenfels, U. Thonemann. 2012. Managers and students as newsvendors. *Management Science* **58**(12) 2225–2233.

Bond, M. 1986. Mutual stereotypes and the facilitation of interaction across cultural lines. *International Journal of Intercultural Relations* **10**(3) 259–276.

Bostian, A., C. Holt, A. Smith. 2008. The newsvendor "pull-to-center effect": Adaptive learning in a laboratory experiment. *Manufacturing & Service Operations Management* **10**(4) 590–608.

Buchan, N., R. Croson. 2004. The boundaries of trust: own and others' actions in the us and china. *Journal of Economic Behavior & Organization* **55**(4) 485–504.

Buchan, N., R. Croson, R. Dawes. 2002. Swift neighbors and persistent strangers: A cross-cultural investigation of trust and reciprocity in social exchange. *American Journal of Sociology* **108**(1) 168–206.

Buchan, N., E. Johnson, R. Croson. 2006. Let's get personal: An international examination of the influence of communication, culture and social distance on other regarding preferences. *Journal of Economic Behavior & Organization* **60**(3) 373–398.

Cachon, G., M. Lariviere. 2001. Contracting to assure supply: How to share demand forecasts in a supply chain. *Management Science* **47**(5) 629–646.

Cai, H., J. Wang. 2006. Overcommunication in strategic information transmission games. *Games & Economic Behavior* **56**(1) 7–36.

Cai, S., M. Jun, Z. Yang. 2010. Implementing supply chain information integration in China: The role of institutional forces and trust. *Journal of Operations Management* **28**(3) 257–268.

Camerer, C. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, Princeton, NJ.

Cason, T., K. Kannan, R. Siebert. 2011. An experimental study of information revelation policies in sequential auctions. *Management Science* **57**(4) 667–688.

Castro, M. 2008. Where are you from? Cultural differences in public good experiments. *Journal of Socio-Economics* **37**(6) 2319–2329.

CDC White Paper. 2011. CDC TradeBeam global trade management overview. October 17, retrieved at `http://www.cdcsoftware.com/en/Resource-Library/Documents/White-Papers-and-Articles/GL-GTM-GB-WP-GTMOverview-PM-09292011`.

Chen, C., M. Peng, P. Saparito. 2002. Individualism, collectivism, and opportunism: A cultural perspective on transaction cost economics. *Journal of Management* **28**(4) 567–583.

Chen, K., M. Kaya, Ö. Özer. 2008. Dual sales channel management with service competition. *Manufacturing & Service Operations Management* **10**(4) 654–675.

Child, J. 1998. Trust and international strategic alliances: The case of Sino-foreign joint ventures. C. Lane, R. Bachmann, eds., *Trust Within and Between Organizations*. Oxford University Press, Oxford, United Kingdom, 241–272.

Child, J., G. Möllering. 2003. Contextual confidence and active trust development in the Chinese business environment. *Organization Science* **14**(1) 69–80.

Chuah, S., R. Hoffmann, M. Jones, G. Williams. 2007. Do cultures clash? Evidence from cross-national ultimatum game experiments. *Journal of Economic Behavior & Organization* **64**(1) 35–48.

Cohen, M., T. Ho, Z. Ren, C. Terwiesch. 2003. Measuring imputed cost in the semiconductor equipment supply chain. *Management Science* **49**(12) 1653–1670.

Cooper, D., J. Kagel, W. Lo, Q. Gu. 1999. Gaming against managers in incentive systems: Experimental results with Chinese students and Chinese managers. *American Economic Review* **89**(4) 781–904.

Crawford, V., J. Sobel. 1982. Strategic information transmission. *Econometrica* **50**(6) 1431–1451.

Croson, R., N. Buchan. 1999. Gender and culture: International experimental evidence from trust games. *American Economic Review* **89**(2) 386–391.

Croson, R., K. Donohue. 2006. Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management Science* **52**(3) 323–336.

Cui, T., J. Raju, Z. Zhang. 2007. Fairness and channel coordination. *Management Science* **53**(8) 1303–1314.

Cui, Y., L. Chen, J. Chen, S. Gavirneni, Q. Wang. 2013. Chinese perspective on newsvendor bias: An exploratory note. *Journal of Operations Management* **31**(1-2) 93–97.

Doney, P., J. Cannon. 1997. An examination of the nature of trust in buyer-seller relationships. *Journal of Marketing* **61**(2) 35–51.

Doney, P., J. Cannon, M. Mullen. 1998. Understanding the influence of national culture on the development of trust. *Academy of Management Review* **23**(3) 601–620.

Dyer, C. 2011. Personal communications with Clint Dyer. Retired Chief Operations Officer, Hitachi Global Storage Technologies.

Eckel, C., R. Wilson. 2004. Is trust a risky decision? *Journal of Economic Behavior & Organization* **55**(4) 447–465.

Egri, C., D. Ralston. 2004. Generation cohorts and personal values: A comparison of China and the United States. *Organization Science* **15**(2) 210–220.

Engle-Warnick, J., R. Slonim. 2006. Learning to trust in indefinitely repeated games. *Games & Economic Behavior* **54**(1) 95–114.

Fan, M. 2006. China's party leadership declares new priority: 'Harmonious society'. *The Washington Post*, October 12. Retrieved at `http://www.washingtonpost.com/wp-dyn/content/article/2006/10/11/AR2006101101610.html`.

Farh, J., A. Tsui, K. Xin, B. Cheng. 1998. The influence of relational demography and *guanxi*: The Chinese case. *Organization Science* **9**(4) 471–488.

Fehr, E., J. List. 2004. The hidden costs and returns of incentives – trust and trustworthiness among CEOs. *Journal of the European Economic Association* **2**(5) 743–771.

Feng, T., L. Keller, X. Zheng. 2011. Decision making in the newsvendor problem: A cross-national laboratory study. *Omega* **39**(1) 41–50.

Fischbacher, U. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* **10**(2) 171–178.

Friedman, D., S. Sunder. 1994. *Experimental Methods: A Primer for Economists*. Cambridge University Press, New York, NY.

Fukuyama, F. 1995. *Trust: The Social Virtues and the Creation of Prosperity*. 28th ed. Free Press, New York, NY.

Gächter, S., B. Herrmann, C. Thöni. 2004. Trust, voluntary cooperation, and socio-economic background: Survey and experimental evidence. *Journal of Economic Behavior & Organization* **55**(4) 505–531.

Gino, F., G. Pisano. 2008. Toward a theory of behavioral operations. *Manufacturing & Service Operations Management* **10**(4) 676–691.

Glaeser, E., D. Laibson, J. Scheinkman, C. Soutter. 2000. Measuring trust. *Quarterly Journal of Economics* **115**(3) 811–846.

Granovetter, M. 1985. Economic action and social structure: The problem of embeddedness. *American Journal of Sociology* **91**(3) 481–510.

Greene, W. 2012. *Econometric Analysis*. Seventh ed. Prentice Hall, Upper Saddle River, NJ.

Gümüş, M. 2013. With or without forecast sharing: Competition and credibility under information asymmetry. Forthcoming in *Production & Operations Management*.

Ha, A., S. Tong. 2008. Contracting and information sharing under supply chain competition. *Management Science* **54**(4) 701–715.

Hagen, J., S. Choe. 1999. Trust in Japanese interfirm relations: Institutional sanctions matter. *Academy of Management Review* **23**(3) 589–600.

Hausman, W. 2011. Private conversation with Warren Hausman and Quantum executive.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath. 2001. In search of homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review* **91**(2) 73–78.

Hewstone, M., M. Rubin, H. Willis. 2002. Intergroup bias. *Annual Review of Psychology* **53**(1) 575–604.

Hinds, P., L. Liu, J. Lyon. 2011. Putting the global in global work: An intercultural lens on the practice of cross-national collaboration. *Academy of Management Annals* **5**(1) 135–188.

Ho, T., K. Weigelt. 2005. Trust building among strangers. *Management Science* **51**(4) 519–530.

Ho, T., J. Zhang. 2008. Designing price contracts for boundedly rational customers: Does the framing of the fixed fee matter? *Management Science* **54**(4) 686–700.

Hofstede, G. 1980. *Culture's Consequences: International Differences in Work-Related Values*. 1st ed. Sage Publications, Inc., Newbury Park, CA.

Hofstede, G. 2001. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. 2nd ed. Sage Publications, Inc., Thousand Oaks, CA.

Holm, H., P. Nystedt. 2005. Intra-generational trust – a semi-experimental study of trust among different generations. *Journal of Economic Behavior & Organization* **58**(3) 403–419.

Hong, K., I. Bohnet. 2007. Status and distrust: The relevance of inequality and betrayal aversion. *Journal of Economic Psychology* **28**(2) 197–213.

Huang, X. 2011. Stereotyping in any form is harmful. *China Daily*, April 13. Retrieved at `http://www.chinadaily.com.cn/opinion/2011-04/13/content_12319727.htm`.

Huff, L., L. Kelley. 2003. Levels of organizational trust in individualist versus collectivist societies: A seven-nation study. *Organization Science* **14**(1) 81–90.

Hyndman, K., S. Kraiselburd, N. Watson. 2013. Aligning capacity decisions in supply chains when demand forecasts are private information: Theory and experiment. *Manufacturing & Service Operations Management* **15**(1) 102–117.

Kalkanci, B., K. Chen, F. Erhun. 2011. Contract complexity and performance under asymmetric demand information: An experimental evaluation. *Management Science* **57**(4) 689–704.

Katok, E., D. Wu. 2009. Contracting in supply chains: A laboratory investigation. *Management Science* **55**(12) 1953–1968.

Knack, S., P. Keefer. 1997. Does social capital have an economic payoff? A cross-country investigation. *Quarterly Journal of Economics* **112**(4) 1251–1288.

Kong, G., S. Rajagopalan, H. Zhang. 2013. Revenue sharing and information leakage in a supply chain. *Management Science* **59**(3) 556–572.

Kreps, D. 1990. Corporate culture and economic theory. J. Alt, K. Shepsle, eds., *Perspectives on Positive Political Economy*. Cambridge University Press, Cambridge, United Kingdom, 90–143.

Kumar, N. 1996. The power of trust in manufacturer-retailer relationships. *Harvard Business Review*, pp. 92–106, November–December.

Kurtuluş, M., S. Ülkü, B. Toktay. 2012. The value of collaborative forecasting in supply chains. *Manufacturing & Service Operations Management* **14**(1) 82–98.

Kuwabara, K., R. Willer, M. Macy, R. Mashima, S. Terai, T. Yamagishi. 2007. Culture, identity, and structure in social exchange: A web-based trust experiment in the United States and Japan. *Social Psychology Quarterly* **70**(4) 461–479.

La Porta, R., F. Lopez de Silanes, A. Shleifer, R. Vishny. 1997. Trust in large organizations. *American Economic Review* **87**(2) 333–338.

Lee, H., V. Padmanabhan, S. Whang. 1997. Information distortion in a supply chain: The bullwhip effect. *Management Science* **43**(4) 546–558.

Lee, H., K. So, C. Tang. 2000. The value of information sharing in a two-level supply chain. *Management Science* **46**(5) 626–643.

Lewicki, R., B. Bunker. 1995. Trust in relationships: A model of development and decline. B. Bunker, J. Rubin, eds., *Conflict, Cooperation and Justice: Essays Inspired by the Work of Morton Deutsch*. Jossey-Bass Publishers, San Francisco, CA, 133–173.

Lim, N., T. Ho. 2007. Designing price contracts for boundedly rational customers: Does the number of blocks matter? *Marketing Science* **26**(3) 312–326.

Lurie, N., J. Swaminathan. 2009. Is timely information always better? The effect of feedback frequency on decision making. *Organizational Behavior & Human Decision Processes* **108**(2) 315–329.

Malhotra, D. 2004. Trust and reciprocity decisions: The differing perspectives of trustors and trusted parties. *Organizational Behavior & Human Decision Processes* **94**(2) 61–73.

Markus, H., S. Kitayama. 1991. Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review* **98**(2) 224–253.

Massey, C., R. Thaler. 2013. The loser's curse: Decision making and market efficiency in the National Football League draft. Forthcoming in *Management Science*.

Matsumoto, D., H. Hwang. 2011. Cooperation and competition in intercultural interactions. *International Journal of Intercultural Relations* **35**(5) 677–685.

McKnight, D., L. Cummings, N. Chervany. 1998. Initial trust formation in new organizational relationships. *Academy of Management Review* **23**(3) 473–490.

Metters, R., X. Zhao, E. Bendoly, B. Jiang, S. Young. 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**(3) 177–185.

Montmarquette, C., J.-L. Rullière, M.-C. Villeval, R. Zeiliger. 2004. Redesigning teams and incentives in a merger: An experiment with managers and students. *Management Science* **50**(10) 1379–1389.

Ockenfels, A., J. Weimann. 1999. Types and patterns: An experimental East-West-German comparison of cooperation and solidarity. *Journal of Public Economics* **71**(2) 275–287.

Özer, Ö., W. Wei. 2006. Strategic commitments for an optimal capacity decision under asymmetric forecast information. *Management Science* **52**(8) 1239–1258.

Özer, Ö., Y. Zheng, K. Chen. 2011. Trust in forecast information sharing. *Management Science* **57**(6) 1111–1137.

Rao, A., J. Pearce, K. Xin. 2005. Governments, reciprocal exchange and trust among business associates. *Journal of International Business Studies* **36**(1) 104–118.

Ren, Z., M. Cohen, T. Ho, C. Terwiesch. 2010. Information sharing in a long-term supply chain relationship: The role of customer review strategy. *Operations Research* **58**(1) 81–93.

Roth, A., V. Prasnikar, M. Okuno-Fujiwara, S. Zamir. 1991. Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study. *American Economic Review* **81**(5) 1068–1095.

Rousseau, D., S. Sitkin, R. Burt, C. Camerer. 1998. Not so different after all: A cross-discipline view of trust. *Academy of Management Review* **23**(3) 393–404.

Schoorman, F., R. Mayer, J. Davis. 2007. An integrative model of organizational trust: Past, present, and future. *Academy of Management Review* **32**(2) 344–354.

Schweitzer, M., G. Cachon. 2000. Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science* **46**(3) 404–420.

Shin, H., T. Tunca. 2010. Do firms invest in forecasting efficiently? The effect of competition on demand forecast investments and supply chain coordination. *Operations Research* **58**(6) 1592–1610.

Singh, R., M. Choo, L. Poh. 1998. In-group bias and fair-mindedness as strategies of self-presentation in intergroup perception. *Personality & Social Psychology Bulletin* **24**(2) 147–162.

Steckel, J., S. Gupta, A. Banerji. 2004. Supply chain decision making: Will shorter cycle times and shared point-of-sales information necessarily help? *Management Science* **50**(4) 458–464.

Sterman, J. 1989. Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science* **35**(3) 321–339.

Su, X. 2008. Bounded rationality in newsvendor models. *Manufacturing & Service Operations Management* **10**(4) 566–589.

Tajfel, H., ed. 2010. *Social Identity and Intergroup Relations*. Reprint ed. Cambridge University Press, Cambridge, UK.

Tajfel, H., J. Turner. 1979. An integrative theory of intergroup conflict. W. Austin, S. Worchel, eds., *The Social Psychology of Intergroup Relations*, chap. 3. Brooks/Cole, Monterey, CA, 33–47.

Takahashi, C., T. Yamagishi, J. Liu, F. Wang, Y. Lin, S. Yu. 2008. The intercultural trust paradigm: Studying joint cultural interaction and social exchange in real time over the Internet. *International Journal of Intercultural Relations* **32**(3) 215–228.

The Economist. 2008. The power of information sharing: Impact on product development. October 22, The Economist Intelligence Unit, retrieved at `http://viewswire.eiu.com/report_dl.asp?mode=fi&fi=1593883744.PDF`.

The Economist. 2012. Boeing – Faster, faster, faster: The planemaker struggles to fulfill a rush of orders. January 28, retrieved at `http://www.economist.com/node/21543555`.

Triandis, H., R. Bontempo, M. Villareal. 1988. Individualism and collectivism: Cross-cultural perspectives on self-ingroup relationships. *Journal of Personality and Social Psychology* **54**(2) 323–338.

Tsui, A., C. Schoonhoven, M. Meyer, C. Lau, G. Milkovich. 2004. Organization and management in the midst of societal transformation: The People's Republic of China. *Organization Science* **15**(2) 133–144.

UBS. 2006. Prices and earnings: A comparison of purchasing power around the globe. Tech. rep.

Valley, K., L. Thompson, R. Gibbons, M. Bazerman. 2002. How communication improves efficiency in bargaining games. *Games & Economic Behavior* **38**(1) 127–155.

Wang, H. 2010. How eBay failed in China. Excerpt from *The Chinese Dream: The rise of the world's largest middle class and what it means to you*. Bestseller Press, Brande, Denmark.

Wicks, A., S. Berman. 2004. The effects of context on trust in firm-stakeholder relationships: The institutional environment, trust creation, and firm performance. *Business Ethics Quarterly* **14**(1) 141–160.

Williamson, O. 1993. Calculativeness, trust, and economic organization. *Journal of Law & Economics* **36**(1) 453–486.

Willinger, M., C. Keser, C. Lohmann, J. Usunier. 2003. A comparison of trust and reciprocity between France and Germany: experimental investigation based on the investment game. *Journal of Economic Psychology* **24**(4) 447–466.

Willnat, L., H. Zhou, X. Hao. 1997. Foreign media exposure and perceptions of Americans in Hong Kong, Shenzhen, and Singapore. *Journalism & Mass Communication Quarterly* **74**(4) 738–756.

World Trade Organization. 2013. *International Trade Statistics 2013*.

World Values Survey. 2009. World Values Survey 1981-2008 Official aggregate v.20090901. Data file, ASEP/JDS, Madrid. Retrieved at `www.worldvaluessurvey.org`.

Xin, K., J. Pearce. 1996. *Guanxi*: Connections as substitutes for formal institutional support. *Academy of Management Journal* **39**(6) 1641–1658.

Yamagishi, T., N. Jin, A. Miller. 1998. In-group bias and culture of collectivism. *Asian Journal of Social Psychology* **1**(3) 315–328.

Yamagishi, T., Y. Makimura, M. Foddy, M. Matsuda, T. Kiyonari, M. Platow. 2005. Comparisons of Australians and Japanese on group-based cooperation. *Asian Journal of Social Psychology* **8**(2) 173–190.

Zak, P., S. Knack. 2001. Trust and growth. *Economic Journal* **111**(470) 295–321.

Zhou, L., M. Hui. 2003. Symbolic value of foreign products in the People's Republic of China. *Journal of International Marketing* **11**(2) 36–58.

## Appendix A: Formulation of the Forecast Sharing Game

Here we present the detailed formulation of the forecast sharing game. Given the retailer's private forecast $\xi$ and the supplier's production quantity $Q$, the supplier's and the retailer's expected profit functions can be characterized as $\Pi^s(\xi,Q)=w\mathbb{E}_\epsilon\min(\xi+\epsilon,Q)-cQ$ and $\Pi^r(\xi,Q)=(r-w)\mathbb{E}_\epsilon\min(\xi+\epsilon,Q)$, respectively. If $\xi$ were known to the supplier, his optimal production quantity would be $Q^*(\xi)=\xi+G^{-1}((w-c)/w)$. This is the optimal newsvendor quantity when the demand follows the distribution of $\epsilon$ and the mean equals $\xi$. In the forecast sharing game, the supplier does not know the exact value of $\xi$ but receives a report $\hat{\xi}$ from the retailer. We note from the expected profit functions that the retailer does not incur any direct cost by reporting, $\hat{\xi}$ is not a binding order, and whether $\hat{\xi}$ equals $\xi$ cannot be perfectly verified even ex post due to the existence of $\epsilon$. Therefore, the retailer's report is costless, nonbinding, and nonverifiable, known as "cheap talk" in the literature (Crawford and Sobel 1982). One approach to solve this game is to apply the solution concept of perfect Bayesian equilibrium. We refer the reader to Özer et al. (2011) for the equilibrium analysis of this game. In the newsvendor task of our forecast sharing experiment, $\xi$ is known to the supplier and the optimal production quantity is indeed $Q^*(\xi)$. Given that $\epsilon$ is uniformly distributed on $[\underline{\epsilon},\bar{\epsilon}]$, we can rewrite the optimal quantity as $Q^*(\xi)=\xi+((w-c)/w)(\bar{\epsilon}-\underline{\epsilon})+\underline{\epsilon}$. As a benchmark, in a centralized supply chain, the optimal production quantity is $Q^c(\xi)=\xi+((r-c)/r)(\bar{\epsilon}-\underline{\epsilon})+\underline{\epsilon}$, and the resulting optimal expected profit is $\Pi^c(\xi,Q^c(\xi))=r\mathbb{E}_\epsilon\min(\xi+\epsilon,Q^c(\xi))-cQ^c(\xi)$.

## Appendix B: Cross-Country Experiment Controls

There are four key factors that need to be controlled for when conducting multi-country experiments: subject pool equivalency, experimenter effect, language effect, and currency effect (Roth et al. 1991). We discuss here how we control for these factors in our experiment. First, all participants were undergraduate students in two leading universities in China and the U.S. with relatively large student population. They majored in the general disciplines of science, engineering, and business. We excluded students from a few social science majors such as history, literature, language, and art. Student participants are commonly used in economic experiments that study fundamental human behavioral factors (such as social preferences) due to the ease of providing proper incentives for careful decisions. In addition, a multitude of studies have proven the value and validity of using student participants to inform our understanding in the behaviors of business professionals (Friedman and Sunder 1994, Chapter 4). In our sample, 41% of the U.S participants and 33% of the Chinese ones were female. 25% of the U.S. participants and 29% of the Chinese ones were majoring in economics or business. The U.S. participants were on average 20 years old (with a standard deviation of 2) and have lived in the U.S. for 17 years on average (with a standard deviation of 7). The Chinese participants were on average 21 years old (with a standard deviation of 1) and have lived in China for 21 years on average (with a standard deviation of 2). The cross-cultural psychology literature shows that residence in a country for over 6 years results in an individual's significant adaptation to the cultures and social norms of the host country (Berry et al. 2006). In our sample, 10 (out of 91) U.S. participants have lived in the U.S. for fewer than 7 years. In addition, 1 (out of 95) Chinese participant has lived in China for only 2 years (she lived in Singapore before that). One may argue that these participants are not good representatives of the two cultures. We verify that all of our results discussed in §5 and §6 remain valid when we remove their data from

the sample. We also perform a stricter robustness check in which we only consider data from the participants who were born in the same country as they have been living in. Again, all of our results discussed in §5 and §6 remain valid. Hence, we continue to use the complete sample for our discussion. Finally, we conducted postexperiment survey to elicit information about the participants' gender, age, school year, major, exposure to game theory, and work experience in a business domain. In our regression analysis, we include control variates for these factors to verify that they are not the driving factors for our conclusions.

To control for experimenter effect, both experimenters ran pilot treatments in the U.S., with no difference found between the two treatments. This procedure to check for possible experimenter effect is widely used in multi-country experimental studies (e.g., Bohnet et al. 2008). To avoid introducing bias through the experimenter, the instructions and practice questions were clearly presented on the computer screen and the experimenter only answered specific questions from individual participants if any. We conducted the sessions using the local language in both countries. To control for language effect, we used back-translation to ensure equivalency between the English and the Chinese descriptions. In particular, we asked two doctoral students who are native Chinese speakers and fluent in English to perform the back-translation. One student translated the materials from English to Chinese. The other student translated the Chinese materials back to English. The experimenters and the two students then compared the original English version with the back-translated version to identify any inconsistencies or confusions caused by the translation. We iterated this process for a few times until all four agreed that the English and Chinese descriptions of the experimental tasks were equivalent. The instructions from both languages are available from the authors upon request.

Finally, we design the payments to achieve equitable strengths of incentives for participants in the two countries. To achieve neutrality of currency units, the profits that participants earned during the experiment were calculated in experimental dollars and converted into actual payments at the end of each session. To calibrate the experimental payments, we consider both the cost of living in the two regions (UBS 2006) and the conventional rate of experimental payments in the two universities. We choose the conversion rates so that on average, U.S. participants earned 30 U.S. dollars and Chinese participants earned 100 Chinese yuan.

## Appendix C: Screenshots of the Experimental Software

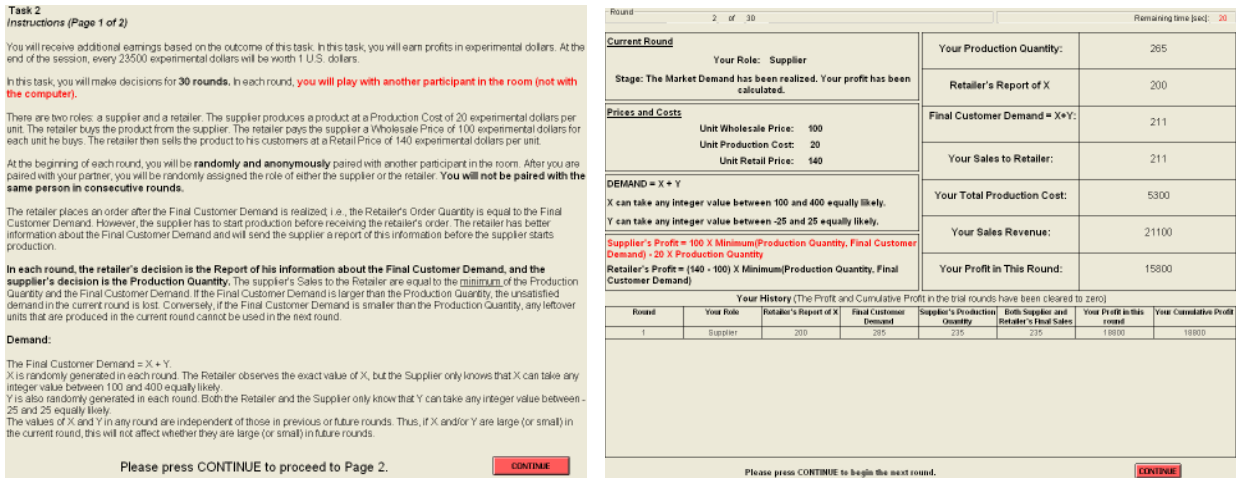
Here we provide two sets of screenshots from the experimental software, one in English and the other in Chinese.

## Appendix D: Regression Analysis of the Newsvendor Task

To examine the treatment effects in the participants' newsvendor decisions in Task 1, we estimate the following random-effects GLMs.

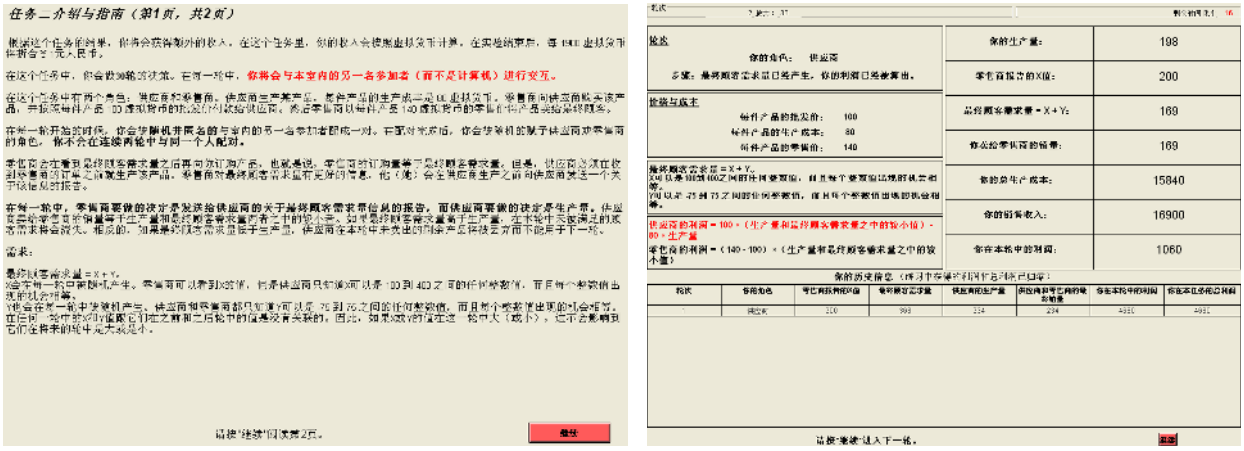
$$\begin{aligned} Q_{1_{it}} &= \text{Intercept} + \lambda_{CN}\cdot CN + \lambda_{C_L}\cdot C_L + \lambda_{CNC_L}\cdot CN\cdot C_L + \lambda_{\xi}\cdot \xi_{it} + \lambda_{C_L t}\cdot C_L\cdot t + \lambda_{C_H t}\cdot (1-C_L)\cdot t \\ &\quad + \delta_i + \varepsilon_{it}, \qquad \text{(A.1)} \\ \left(\frac{Q_1-\xi}{Q_1^*(\xi)-\xi}\right)_{it} &= \text{Intercept} + \lambda_{CN}\cdot CN + \lambda_{C_L}\cdot C_L + \lambda_{CNC_L}\cdot CN\cdot C_L + \lambda_t\cdot t + \delta_i + \varepsilon_{it}. \qquad \text{(A.2)} \end{aligned}$$

The variables $CN$ and $C_L$ are the dummy variables for the treatment conditions: $CN = 1$ if the participant is from China and 0 otherwise; $C_L = 1$ for a low-cost treatment and 0 otherwise. $\xi_{it}$ is included in Equation




(a) Sample Instructions Page (b) Sample Results Page

**Figure A.1 Sample Screenshots of the Experimental Software in English**




(a) Sample Instructions Page (b) Sample Results Page

**Figure A.2 Sample Screenshots of the Experimental Software in Chinese**

(A.1) to control for the dependency between the participants' production decisions and the actual forecasts. $t$ is included in both equations to capture possible time effects in the participants' decisions. We interact the cost condition with $t$ in Equation (A.1) to capture the possible different directions of time effects in $Q_1$ due to learning. That is, if individuals learn to make better newsvendor decisions, they will produce smaller (larger) quantities under a high (low) cost over time (e.g., Bolton and Katok 2008). $\delta_i$ represents the individual-specific error and $\varepsilon_{it}$ captures the independent errors across decisions.

Table A.1 summarizes the estimates and significance results for the coefficient terms that are relevant to our comparisons. The first column of the table explains the comparison captured by each term.

## Appendix E: Regression Analysis of the Forecast Sharing Task

Here we present the detailed results of the regression analysis discussed in §5.2 and §5.3. The coefficient estimates and comparison results for Equations (1) and (2) are reproduced in Columns 2 and 5 of Tables

**Table A.1 Comparison of Production Quantities and Adjustment Scores in the Newsvendor Task**

| | | Value (standard error) | |
|---|---|---|---|
| **Comparison** | **Coefficients** | $Q_1$ | $(Q_1-\xi)/(Q_1^*(\xi)-\xi)$ |
| China vs. U.S. under a high cost | $\lambda_{CN}$ | -0.50 (4.82) | 0.01 (0.11) |
| China vs. U.S. under a low cost | $\lambda_{CN}+\lambda_{CNC_L}$ | -4.55 (4.79) | -0.10 (0.11) |
| U.S., low vs. high cost | $\lambda_{C_L}$ | 32.18 (5.01)$^{**}$ | 0.44 (0.11)$^{**}$ |
| China, low vs. high cost | $\lambda_{C_L}+\lambda_{CNC_L}$ | 28.14 (4.94)$^{**}$ | 0.33 (0.11)$^{**}$ |

$**: p<0.01$; $p$ values are derived from $t$ tests.

**Table A.2 Tobit Regressions on Report $\hat{\xi}$**

**Regression estimates**

| | Value (standard error) | | | | | |
|---|---|---|---|---|---|---|
| **Variable** | **High cost** | | | **Low cost** | | |
| Intercept | 58.14 (10.48)$^{**}$ | 58.07 (10.23)$^{**}$ | 105.17 (54.44)$^{*}$ | 75.85 (9.27)$^{**}$ | 77.80 (9.08)$^{**}$ | 128.51 (79.25)$^{\dagger}$ |
| *CN* | 20.20 (13.76)$^{\dagger}$ | 19.96 (13.27)$^{\dagger}$ | 20.89 (14.17)$^{\dagger}$ | 30.48 (12.14)$^{**}$ | 29.11 (11.75)$^{**}$ | 31.31 (12.67)$^{**}$ |
| *CNP* | 5.06 (13.76) | 3.49 (13.28) | 13.86 (13.34) | -18.44 (12.41)$^{\dagger}$ | -19.26 (12.01)$^{\dagger}$ | -16.48 (12.48)$^{\dagger}$ |
| $CN\cdot CNP$ | 0.51 (19.32) | -1.20 (18.65) | -9.47 (18.66) | 19.16 (17.43) | 15.24 (16.93) | 23.88 (17.45)$^{\dagger}$ |
| $t$ | 1.37 (0.28)$^{**}$ | 0.99 (0.32)$^{**}$ | 1.37 (0.29)$^{**}$ | 1.14 (0.30)$^{**}$ | 0.81 (0.33)$^{**}$ | 1.14 (0.30)$^{**}$ |
| $\xi$ | 0.82 (0.01)$^{**}$ | 0.82 (0.01)$^{**}$ | 0.82 (0.01)$^{**}$ | 0.79 (0.02)$^{**}$ | 0.79 (0.02)$^{**}$ | 0.79 (0.02)$^{**}$ |
| $(Q_2-\hat{\xi})_{i,t-1}$ | – | -0.08 (0.03)$^{**}$ | – | – | -0.09 (0.03)$^{**}$ | – |
| Male | – | – | 2.96 (9.92) | – | – | 15.60 (9.62)$^{\dagger}$ |
| Age | – | – | -2.11 (2.71) | – | – | -3.10 (3.94) |
| Economics Major | – | – | 8.80 (13.19) | – | – | -7.87 (11.28) |
| Game Theory | – | – | -11.04 (9.74) | – | – | 3.40 (9.56) |
| Work Years | – | – | -8.78 (5.43)$^{\dagger}$ | – | – | -2.39 (5.14) |
| **Comparison & coefficient**$^{\ddagger}$ | | | | | | |
| $S_{CN}R_{CN}-S_{US}R_{US}$: $\lambda_{CN}+\lambda_{CNP}+\lambda_{CN\cdot CNP}$ | 25.77 (13.76)$^{*}$ | 22.25 (13.30)$^{*}$ | 25.28 (13.87)$^{*}$ | 31.20 (12.14)$^{**}$ | 25.09 (11.88)$^{*}$ | 38.71 (14.04)$^{**}$ |
| $S_{US}R_{CN}-S_{US}R_{US}$: $\lambda_{CN}$ | 20.20 (13.76)$^{\dagger}$ | 19.96 (13.27)$^{\dagger}$ | 20.89 (14.17)$^{\dagger}$ | 30.48 (12.14)$^{**}$ | 29.11 (11.75)$^{**}$ | 31.31 (12.67)$^{**}$ |
| $S_{CN}R_{CN}-S_{CN}R_{US}$: $\lambda_{CN}+\lambda_{CN\cdot CNP}$ | 20.71 (13.44)$^{\dagger}$ | 18.76 (12.96)$^{\dagger}$ | 11.42 (14.46) | 49.64 (12.41)$^{**}$ | 44.35 (12.13)$^{**}$ | 55.19 (14.86)$^{**}$ |
| $S_{CN}R_{US}-S_{US}R_{US}$: $\lambda_{CNP}$ | 5.06 (13.76) | 3.49 (13.28) | 13.86 (13.34) | -18.44 (12.41)$^{\dagger}$ | -19.26 (12.01)$^{\dagger}$ | -16.48 (12.48)$^{\dagger}$ |
| $S_{CN}R_{CN}-S_{US}R_{CN}$: $\lambda_{CNP}+\lambda_{CN\cdot CNP}$ | 5.57 (13.44) | 2.29 (12.98) | 4.39 (12.43) | 0.72 (12.14) | -4.02 (11.81) | 7.40 (12.14) |

Notes. "–" means the corresponding variable is not present in the regression.

‡: Values show changes in the dependent variable between the two treatments.

$**: p<0.01$; $*: p<0.05$; $\dagger: p<0.1$; $p$ values are derived from $t$ tests.

A.2 and A.3 to facilitate comparison. The remaining columns in these tables present the results for the additional models that control for experience, gender, age, majoring in economics or business related disciplines, exposure to game theory, and years of work experience in a business domain.

Regarding experience effect, a retailer's decision in the current round may be affected by how much the previous supplier trusted his/her report. The level of trust from the previous supplier is captured in the supplier's production adjustment $(Q_2-\hat{\xi})$. The lower this value, the less trust the retailer received from the previous supplier, all else being equal. Thus, we include $(Q_2-\hat{\xi})_{i,t-1}$ as an independent variable in Equation (1), where $t$ is the current round index. Similarly, a supplier's decision in the current round may be affected by his/her evaluation of how trustworthy the previous retailer was. This evaluation can be captured by the difference between the report and the realized demand $(\hat{\xi}-D)$. The higher this value, the less trustworthy the supplier regards the previous retailer to be, all else being equal. Hence, we include the term $(\hat{\xi}-D)_{i,t-1}$

**Table A.3 Regressions on Production Adjustment $Q_2 - \hat{\xi}$**

**Regression estimates**

| | Value (standard error) | | | | | |
|---|---|---|---|---|---|---|
| **Variable** | **High cost** | | | **Low cost** | | |
| Intercept | -0.167 (8.01) | -1.74 (8.10) | 9.27 (52.29) | 14.45 (9.85)$^{\dagger}$ | 14.04 (10.11)$^{\dagger}$ | -19.04 (35.79) |
| *CN* | -17.28 (10.14)$^{\dagger}$ | -16.98 (10.08)$^{*}$ | -15.16 (10.81)$^{\dagger}$ | 1.43 (10.15) | -0.02 (10.33) | -0.60 (9.36) |
| *CNP* | -10.27 (10.27) | -7.55 (10.22) | -10.18 (10.62) | -14.94 (9.92)$^{\dagger}$ | -12.00 (10.12) | -12.60 (9.76)$^{\dagger}$ |
| $CN \cdot CNP$ | -8.54 (14.08) | -7.77 (14.01) | -8.72 (14.12) | -46.37 (14.23)$^{**}$ | -46.17 (14.50)$^{**}$ | -45.68 (13.66)$^{**}$ |
| $t$ | -3.48 (0.34)$^{**}$ | -2.74 (0.38)$^{**}$ | -3.48 (0.34)$^{**}$ | -3.32 (0.30)$^{**}$ | -2.77 (0.34)$^{**}$ | -3.32 (0.30)$^{**}$ |
| $(\overline{Q_1 - \xi})$ | 0.55 (0.20)$^{**}$ | 0.51 (0.20)$^{**}$ | 0.56 (0.21)$^{**}$ | 0.61 (0.23)$^{**}$ | 0.61 (0.24)$^{**}$ | 0.72 (0.22)$^{**}$ |
| $(\hat{\xi} - D)_{i,t-1}$ | – | -0.18 (0.03)$^{**}$ | – | – | -0.12 (0.02)$^{**}$ | – |
| Male | – | – | 3.28 (7.73) | – | – | -10.44 (7.62) |
| Age | – | – | -0.82 (2.47) | – | – | 2.00 (1.73) |
| Economics Major | – | – | 1.79 (8.12) | – | – | 11.34 (8.43) |
| Game Theory | – | – | 5.95 (7.99) | – | – | -15.14 (7.58)$^{*}$ |
| Work Years | – | – | 3.40 (4.36) | – | – | -5.00 (2.41)$^{*}$ |
| **Comparison & coefficient**$^{\ddagger}$ | | | | | | |
| $S_{CN}R_{CN} - S_{US}R_{US}$: $\lambda_{CN} + \lambda_{CNP} + \lambda_{CN \cdot CNP}$ | -36.09 (10.11)$^{**}$ | -32.30 (10.08)$^{**}$ | -34.06 (10.46)$^{**}$ | -59.88 (9.92)$^{**}$ | -58.19 (10.12)$^{**}$ | -58.88 (9.39)$^{**}$ |
| $S_{CN}R_{US} - S_{US}R_{US}$: $\lambda_{CN}$ | -17.28 (10.14)$^{*}$ | -16.98 (10.08)$^{*}$ | -15.16 (10.81)$^{\dagger}$ | 1.43 (10.15) | -0.02 (10.33) | -0.60 (9.36) |
| $S_{CN}R_{CN} - S_{US}R_{CN}$: $\lambda_{CN} + \lambda_{CN \cdot CNP}$ | -25.82 (9.91)$^{**}$ | -24.75 (9.85)$^{**}$ | -23.88 (11.63)$^{*}$ | -44.94 (9.88)$^{**}$ | -46.19 (10.05)$^{**}$ | -46.28 (10.74)$^{**}$ |
| $S_{US}R_{CN} - S_{US}R_{US}$: $\lambda_{CNP}$ | -10.27 (10.27) | -7.55 (10.22) | -10.18 (10.62) | -14.94 (9.92)$^{\dagger}$ | -12.00 (10.12) | -12.60 (9.76)$^{\dagger}$ |
| $S_{CN}R_{CN} - S_{CN}R_{US}$: $\lambda_{CNP} + \lambda_{CN \cdot CNP}$ | -18.81 (9.64)$^{*}$ | -15.32 (9.60)$^{\dagger}$ | -18.90 (9.51)$^{*}$ | -61.31 (10.11)$^{**}$ | -58.17 (10.35)$^{**}$ | -58.28 (9.20)$^{**}$ |

Notes. "–" means the corresponding variable is not present in the regression.

‡: Values show changes in the dependent variable between the two treatments.

$**: p < 0.01$; $*: p < 0.05$; $\dagger: p < 0.1$; $p$ values are derived from $t$ tests.

as an independent variable in Equation (2). Columns 3 and 6 of Tables A.2 and A.3 show that the coefficients for the experience terms are all significantly negative. These results suggest that the retailers inflate more when $(Q_2 - \hat{\xi})$ from the previous round is lower, i.e., when the previous supplier is less trusting. Conversely, the suppliers produce less when $(\hat{\xi} - D)$ from the previous round is higher, i.e., when the previous retailer is less trustworthy. Also note that when these experience terms are controlled for in the regression models, the magnitude of the coefficient estimates for $t$ is reduced. Thus, past experience contributes to the observed time effects in participants' decisions. We also test experience effects by controlling for the average values of the above two terms from all past rounds, as well as including interactions between treatment dummies and the experience variables. These additional models yield similar results as the ones presented here and are thus omitted.

Finally, we emphasize that all of the observations discussed in §5.2 remain valid under these additional models, suggesting that our observations capture systematic patterns in the participants' forecast sharing behavior.

## Appendix F: Regression Analysis of the Repeated-Interaction Treatments

We first present the detailed regression models and parameter estimates for testing Hypothesis 5. We use random-effects GLMs to analyze treatment effects for three dependent variables: the retailers' reports $\hat{\xi}$

(Tobit model is used), the suppliers' production adjustment $(Q_2 - \hat{\xi})$, and supply chain efficiency as defined in §5.2.3. The regression models are summarized below.

$$\begin{aligned}
\hat{\xi}^*_{it} &= \text{Intercept} + \lambda_{CN} \cdot CN + \lambda_{CNP} \cdot CNP + \lambda_{CN\,CNP} \cdot CN \cdot CNP + \lambda_{REP} \cdot REP + \lambda_{CN\,REP} \cdot CN \cdot REP \\
&\quad + \lambda_{CNP\,REP} \cdot CNP \cdot REP + \lambda_{CN\,CNP\,REP} \cdot CN \cdot CNP \cdot REP + \lambda_{\xi} \cdot \xi_{it} + \lambda_t \cdot t + \delta_i + \varepsilon_{it}, \qquad \text{(A.3)} \\
(Q_2 - \hat{\xi})_{it} &= \text{Intercept} + \lambda_{CN} \cdot CN + \lambda_{CNP} \cdot CNP + \lambda_{CN\,CNP} \cdot CN \cdot CNP + + \lambda_{REP} \cdot REP + \lambda_{CN\,REP} \cdot CN \cdot REP \\
&\quad + \lambda_{CNP\,REP} \cdot CNP \cdot REP + \lambda_{CN\,CNP\,REP} \cdot CN \cdot CNP \cdot REP + \lambda_{Q1adj} \cdot (\overline{Q_1 - \xi})_i + \lambda_t \cdot t + \delta_i + \varepsilon_{it}, \qquad \text{(A.4)} \\
E_{it} &= \text{Intercept} + \lambda_{CN} \cdot CN + \lambda_{CNP} \cdot CNP + \lambda_{CN\,CNP} \cdot CN \cdot CNP + + \lambda_{REP} \cdot REP + \lambda_{CN\,REP} \cdot CN \cdot REP \\
&\quad + \lambda_{CNP\,REP} \cdot CNP \cdot REP + \lambda_{CN\,CNP\,REP} \cdot CN \cdot CNP \cdot REP + \lambda_{\xi} \cdot \xi_{it} + \lambda_t \cdot t + \delta_i + \varepsilon_{it}. \qquad \text{(A.5)}
\end{aligned}$$

The dummy variable *REP* indicates repeated interactions: $REP = 1$ if the data is from a repeated-interaction treatment and 0 otherwise. Columns 2–4 of Table A.4 summarize the regression estimates.

To study the participants' behavioral dynamics in the repeated-interaction treatments, we estimate two additional regression models whose estimates are reported in Columns 5–6 of Table A.4. These regression models differ from Equations (A.3) and (A.4) in that we include the interaction terms between the treatment dummies and the round as well as experience variables (see the last 15 rows in Table A.4). As before, we use $(Q_2 - \hat{\xi})_{i,t-1}$ to capture retailer $i$'s experience and $(\hat{\xi} - D)_{i,t-1}$ to capture supplier $i$'s experience.

## Appendix G: Individual Heterogeneity in the Repeated-Interaction Treatments

One puzzling observation in the current data is that the U.S. supply chain does not seem to react strongly to the presence of repeated interactions (see Özer et al. 2011 for a contrasting result). To understand why, we dig deeper into individual heterogeneity for the U.S. participants. We notice that there are four retailers in the $S_{US}R_{US}$REP treatment whose forecast inflation has a standard deviation that is more than twice higher than the maximum standard deviation among the other retailers in all treatments. We conjecture that these four retailers have somewhat unique behavioral patterns and hence may impact the efficiency of the supply chains involving them. To test this conjecture, we divide the twelve pairs of participants in $S_{US}R_{US}$REP into two groups: the low-variance and high-variance groups. We then re-estimate Equation (A.5) using either only the low-variance or only the high-variance group data for $S_{US}R_{US}$REP (data used for the other treatments remain unchanged). The results show that the efficiency of a U.S. supply chain increases by 8.55% on average ($p = 0.01$) under repeated interactions with only the low-variance group data, whereas it decreases by 13.58% ($p = 0.001$) with only the high-variance group data. Thus in aggregate (i.e., when data from both groups are used), we do not observe a significant change. When examining the suppliers' decisions, we observe that the highly variable behavior of retailers in the high-variance group leads to a stronger decline of trust by the paired suppliers, which hurts the supply chain efficiency. An interesting question thus arises as to whether individual heterogeneity induces more variance in the performance of a U.S. supply chain (versus a Chinese one), possibly due to the immigrant and individualistic culture of the country. Future studies can help to answer this question.

**Table A.4 Regression Estimates for Repeated-Interaction Treatments**

| | Estimates for Equations (A.3) – (A.5) Value (standard error) | | | Estimates for Behavioral Dynamics Value (standard error)‡ | |
|---|---|---|---|---|---|
| **Variable** | $\hat{\xi}$ | $Q_2-\hat{\xi}$ | **Efficiency (%)** | $\hat{\xi}$ | $Q_2-\hat{\xi}$ |
| Intercept | 47.75 (8.61)** | -3.18 (8.47) | 81.29 (2.61)** | 45.36 (9.37)** | -9.77 (9.79) |
| *CN* | 20.50 (11.44)* | -16.26 (11.16)† | -4.67 (2.98)† | 18.17 (12.54)† | -13.53 (13.18) |
| *CNP* | 5.36 (11.44) | -12.25 (11.23) | -4.04 (2.98)† | 13.38 (12.63) | 4.98 (13.24) |
| *CN* · *CNP* | 0.20 (15.99) | -7.09 (15.42) | -1.74 (4.12) | -7.90 (17.56) | -3.14 (18.23) |
| *REP* | -11.63 (11.40) | -8.18 (11.10) | -3.30 (2.98) | -4.02 (12.58) | -10.96 (13.16) |
| *CN* · *REP* | -7.29 (15.99) | 26.87 (15.42)* | 13.26 (4.12)** | -12.19 (17.50) | 26.01 (18.25)† |
| *CNP* · *REP* | -4.24 (15.99) | 17.84 (15.44) | 8.88 (4.12)* | -25.12 (17.57) | 11.03 (18.22) |
| *CN* · *CNP* · *REP* | 0.32 (22.48) | -19.15 (21.51) | -7.25 (5.77) | 19.81 (24.75) | -36.59 (25.46) |
| $\xi$ | 0.87 (0.01)** | – | 0.03 (0.01)** | 0.87 (0.01)** | – |
| $\overline{Q_1-\xi}$ | – | 0.76 (0.16)** | – | – | 0.67 (0.15)** |
| $t$ | 0.95 (0.20)** | -2.81 (0.23)** | 0.11 (0.12) | 1.11 (0.68) | -1.79 (0.80)* |
| *CN* · $t$ | – | – | – | -0.11 (0.93) | 0.10 (1.08) |
| *CNP* · $t$ | – | – | – | -1.78 (0.89)* | -1.77 (1.10) |
| *CN* · *CNP* · $t$ | – | – | – | 2.59 (1.22)* | -0.52 (1.50) |
| *REP* · $t$ | – | – | – | -0.29 (0.89) | 0.18 (1.07) |
| *CN* · *REP* · $t$ | – | – | – | 0.04 (1.23) | 0.06 (1.50) |
| *CNP* · *REP* · $t$ | – | – | – | 2.04 (1.20)† | 1.28 (1.50) |
| *CN* · *CNP* · *REP* · $t$ | – | – | – | -2.76 (1.67)† | 1.78 (2.09) |
| Exp | – | – | – | -0.02 (0.07) | -0.07 (0.06) |
| *CN*·Exp | – | – | – | -0.09 (0.10) | -0.16 (0.08)* |
| *CNP*·Exp | – | – | – | -0.10 (0.09) | -0.08 (0.08) |
| *CN* · *CNP*·Exp | – | – | – | 0.12 (0.12) | 0.09 (0.11) |
| *REP*·Exp | – | – | – | 0.18 (0.09)* | -0.01 (0.07) |
| *CN* · *REP*·Exp | – | – | – | -0.16 (0.13) | -0.07 (0.11) |
| *CNP* · *REP*·Exp | – | – | – | -0.34 (0.14)* | -0.14 (0.11) |
| *CN* · *CNP* · *REP*·Exp | – | – | – | 0.40 (0.19)* | 0.22 (0.16) |

Note. "–" means the variable is not present in the corresponding model.

‡: Notation "Exp" in the last 8 rows stands for $(Q_2-\hat{\xi})_{i,t-1}$ and $(\hat{\xi}-D)_{i,t-1}$ for the regression on $\hat{\xi}$ and $(Q_2-\hat{\xi})$, respectively.

$**: p<0.01$; $*: p<0.05$; $\dagger: p<0.1$; $p$ values are derived from $t$ tests.