DRAFT
DO NOT DISTRIBUTE without explicit authorization

Will You Automate Me? Delegating B2B Negotiation to Human or Bot

Klaus Garrido Tenorio,[a] Michael Becker-Peth,[a] Jelle de Vries,[b]
[a]Department of Technology & Operations Management, Rotterdam School of Management, Erasmus University, garridotenorio@rsm.nl, m.beckerpeth@rsm.nl; [b]Department of Information Systems and Operations Management, Tilburg School of Economics and Management, Tilburg University, jvries@tilburguniversity.edu

Abstract. Firms increasingly delegate procurement and sales negotiations to autonomous AI agents, yet little is known about how managers decide whether to send a bot or a human to negotiate. We frame this decision as a choice between a general investment, which strengthens a firm's outside options, and a costly relationship-specific investment, which can signal trust to a counterpart. We conduct a preregistered experiment in a two-tier supply chain in which a retailer and a supplier manager sequentially delegate a wholesale price and quantity negotiation either to a dedicated human agent (relationship-specific investment) or to a general agent, which is human in the control and an AI bot in the treatment. We vary the product class so that the risky relationship-specific investment yields a higher, equal, or lower expected payoff than the certain general investment. We find modest evidence of AI delegation avoidance: when the general investment is an AI bot, managers are more likely to choose the human relationship-specific agent. This avoidance is contingent on the economics of the exchange. It is most pronounced in the product class with high transaction cost, where relationship building pays off in expectation, and not detectable in the trans-actional class, where the general investment is superior. A post-hoc analysis of the subsequent negotiations shows that agreements between two human agents order fewer units than the Nash bargaining solution and are less efficient than agreements involving a bot. The expected transaction-cost savings of a reciprocated human relationship do not compensate for this loss in deal efficiency.

Funding: This research was supported by Erasmus Research Institute of Management (ERIM), by a research grant from Negotiation and Team Resources (NTR), and the Dutch Research Council (NWO) with a Small Compute Grant EINF-10333.

Key words: Buyer-Supplier Negotiation, Human-Bot Interaction, Delegation, Lab Experiment

1. Introduction

Business to Business (B2B) supply chain interactions are increasingly being automated. In procurement, the Artificial Intelligence (AI) label in chatbot systems deployed for Request-for-Quote (RFQ) leads to lower prices than human agents (Cui et al. 2022). Large corporations like Walmart, that started piloting automating negotiations with tail-end suppliers back in 2021 rely on Information Technology (IT) providers like Pactum rather than in-house development (van Hoek et al. 2022). Generally, automated negotiations focus on routine low-value contracts with few terms (Falcão Filho 2024, Revilla and Saenz 2026). However, Pactum's technology recently dealt with a high-stakes supplier relationship valued at approximately $530 million (Forbes 2026), highlighting the unprecedented scalability of automated bargaining. As Jessen-Klixbüll, Head of Global Procurement at Maersk, observed regarding this operational shift (van Hoek and Lacity 2023):

> *"The ability to take on a massive amount of negotiations simultaneously and to be able to scale that across a number of scenarios is of incredible value."*

While some supply chain B2B interactions shift toward Bot-to-Bot exchanges (Guida et al. 2023), human negotiators expecting to deal with a fellow human may be interacting with a bot counterpart. Such Human-Bot interactions may forgo the relationship benefits that partner-specific investments like dedicated human resources can generate (Beer et al. 2018). Consequently, as managers gain the option to delegate B2B negotiations to artificial agents, they face a trade-off between the certain cost savings of artificial agent and the uncertain relationship benefits of a costlier human agent. It remains unclear which they will choose.

By investing in a specific relationship through the dedication of human resources to negotiate with a counterpart, a firm signals inherent trustworthiness, which can encourage the counterpart to reciprocate through higher effort and the provision of non-contractible quality (Beer et al. 2018). In contrast, automated bot deployment constitutes a general investment that increases the buyer's bargaining power by improving its outside options (Beer et al. 2018). For instance, empirical evidence from Alibaba's supplier base demonstrates that suppliers quote lower prices to chatbot buyers that disclose their automated "smart" nature compared to human buyers or unlabelled bots (Cui et al. 2022). The underlying rationale is that suppliers perceive the analytical capabilities of AI as a signal of strong outside options, assuming an AI buyer can readily compare and contact hundreds of alternative suppliers (Cui et al. 2022). Such general investments secure better transactional deals for the buyer but forgo the relational benefits that human contact generates. While managers along the supply chain increasingly face artificial agents as negotiation counterparts, existing literature has

focused primarily on examining the standalone bargaining behavior of large language models (Vaccaro et al. 2025, Kirshner et al. 2026). Although recent work investigates the behavioral responses of human negotiators when choosing between artificial and human counterparts (Garrido Tenorio et al. 2027, Herold et al. 2025, Chen and Huang 2026), empirical evidence regarding the decision to delegate a buyer-supplier negotiation to artificial agents rather than humans remains limited. While delegation has been extensively examined in objective task domains(Fügener et al. 2022, Bockstedt and Buckman 2025), extending these insights to subjective, interactive environments requires looking closer at interpersonal dynamics. Evidence from business-to-consumer (B2C) customer service contexts demonstrates that end-users consistently prefer human agents over machines when exchanges involve subjective and interactive elements (Castelo et al. 2019, Dai et al. 2026). To examine how these dynamics translate to B2B procurement, we investigate the research question:

"How does the nature of the delegated agent (Human vs. AI) shape the decision to make a relation-specific or general investment in B2B negotiations?"

To address this research question, we design and conduct a pre-registered, two-stage laboratory experiment embedded in a two-tier supply chain bargaining setting (Davis and Hyndman 2019). In the first stage, retailer and supplier managers sequentially delegate an upcoming wholesale price and quantity negotiation. Operating under a Stackelberg dynamic, the first-mover manager selects a representative, and the second-mover manager observes this choice before making their own delegation decision. Managers choose between a costly relationship-specific investment (delegating to Agent type-A) and a lower-cost general investment (delegating to Agent type-B). Across conditions, Agent type-A is always human, whereas our experimental treatment manipulates whether Agent type-B is an artificial or a human agent. While delegating to Agent type-B guarantees lower upfront delegation costs, mutual delegation to Agent A can establish a collaborative relationship that halves ex-post transaction costs (Beer et al. 2018). To control for the economic trade-off between guaranteed cost efficiency and uncertain relational gains, we vary ex-post transaction costs across three product classes regimes (High-, Mid-, and Low-Transaction Cost) such that the relationship-specific investment yields lower, equal, or higher expected costs than the general investment. In the second stage, the delegated agents execute the unstructured negotiation across the resulting Human-to-Human, Human-to-Bot, or Bot-to-Bot dyads, whose realized agreements determine final company and managerial payoffs.

We find evidence, albeit modest, of AI delegation avoidance. When the general investment entails delegating the negotiation to a bot rather than a human, the odds that managers choose the

relationship-specific human agent increase by approximately 47% ($p = 0.073$, one-sided), and the effect persists after controlling for the product class. This avoidance is contingent on the economics of the exchange. In the strategic product class, where the risky relationship-specific investment offers the higher expected return, the presence of AI raises the share of managers choosing the human relationship by 16.4%, from 64.1% to 80.5%. In the low-transaction cost product class, where the certain general investment is the superior alternative, managers do not penalize the bot. A post-hoc analysis of the subsequent negotiations reveals a further pattern. Agreements between two human negotiators order fewer units than the Nash bargaining solution and lose more expected joint deal profit relative to that benchmark than agreements involving a bot. In our setting, the expected transaction-cost savings of a reciprocated human relationship do not offset this loss in deal efficiency.

Our study contributes to the literature on human–AI delegation, which has examined delegation almost exclusively in structured, objective tasks such as image classification and forecasting (Fügener et al. 2022, Bockstedt and Buckman 2025). We show that a preference for human agents extends to the unstructured, relational domain of B2B negotiations, where the principal delegates the interaction but retains its economic consequences. By framing the choice of negotiator through the distinction between general and relationship-specific precontractual investments (Beer et al. 2018), we connect this behavioral preference to the economics of supply chain relationships and show that AI delegation avoidance depends on the expected returns of relationship building rather than operating as an unconditional bias. Finally, we add to the behavioral bargaining literature, which documents that human negotiators order inefficiently low quantities against both human (Davis and Hyndman 2019) and bot counterparts (Chen and Huang 2026). We show that a bot anchored at the Nash bargaining solution brings deals with human counterpart to higher supply-chain efficiency than human-human negotiations.

For managers, our findings carry two implications. First, offering the option to delegate negotiations to AI does not appear to degrade delegation decisions: managers reserve human negotiators for strategic counterparts, where relationship building pays off in expectation, and accept AI delegation where the general investment is economically superior. Firms can therefore introduce AI negotiation agents while relying on managerial discretion to decide which relationships merit a human touch. Second, the value of a human relationship must be weighed against the efficiency of the deals it produces. Even when managers reciprocate a counterpart's relationship-specific investment, as our results suggest they are more inclined to do when the alternative is a bot, two human negotiators tend

to reach less efficient agreements than human-bot or bot-bot negotiations. Procurement managers who choose human delegation for relational reasons should therefore recognize that the relational benefits must be large enough to compensate for the loss in deal efficiency.

2. Literature Review & Hypothesis Development

Our study builds on two streams of literature. The first, reviewed in Section 2.1, examines how humans interact with AI negotiators and how they delegate tasks to algorithms. It suggests that principals may avoid delegating subjective, relational exchanges to machines. The second, reviewed in Section 2.2, comes from supply chain management and distinguishes general from relationship-specific precontractual investments. It allows us to assess the economic consequences of the delegation choice and to derive our hypothesis.

2.1. Human-Bot Interaction

AI negotiators are moving from pilot projects to high-stakes procurement relationships, and human negotiators increasingly face artificial counterparts (Revilla and Saenz 2026). This subsection first reviews industry evidence on the deployment of negotiation bots. It then summarizes behavioral research on how humans respond to algorithmic advisors and bargaining counterparts. Finally, it turns to the open question of how managers decide whether to delegate an interpersonal negotiation to a machine.

AI chatbots have demonstrated substantial value in procurement interactions with suppliers. Early adopters of AI negotiators report substantial efficiency gains. Specifically, automating negotiations for annual spending of $7 billion has been shown to extend payment terms to an average of 35 days and achieve cost savings of at least 5% (van Hoek and Lacity 2023, van Hoek 2024). Historically, large buyers limited AI negotiations to tail spend, those transactional categories with non-critical terms where buyer involvement is minimal and relational depth is inconsequential (Falcão Filho 2024, Borsani et al. 2026). Demonstrating a distinct shift away from this constraint, an AI negotiation conducted by Pactum's software managed a high-stakes supplier relationship valued at approximately $530 million (Forbes 2026). Although technology providers initially designed negotiation bots to interact exclusively with human counterparts, the opposing party may also deploy an artificial agent. IT experts interviewed by Guida et al. (2023) anticipated this rise in bot-to-bot negotiation, suggesting that buyer-supplier negotiations will transition to be handled by bots that maximize their assigned objective functions. Yet, whereas media coverage claims symmetric, fully autonomous bot-to-bot exchanges (McManus 2023), peer-reviewed outlets suggest

an asymmetric adoption pattern. In practice, large enterprise buyers primarily deploy automated bots to negotiate with human representatives of tail-end suppliers (van Hoek et al. 2022, Revilla and Saenz 2026). Because human negotiators increasingly confront artificial counterparts in the field, understanding behavioral responses to algorithmic bargaining agents represents a critical research priority (Kirshner et al. 2026).

A rapidly expanding literature examines the behavior of humans interacting with AI and of AI itself across two primary dimensions: whether large language models (LLMs) replicate human behavior and how human decision-makers respond to algorithmic counterparts and their outputs (Chen et al. 2025, Caro et al. 2026). Regarding machine behavior, recent experimental studies investigate whether LLMs exhibit human cognitive biases and bargaining heuristics (Chen et al. 2025, Vaccaro et al. 2025, Kirshner et al. 2026). Specifically, Kirshner et al. (2026) demonstrate that in bot-to-bot supply chain negotiations, . Regarding human reception, behavioral operations research extensively documents how individuals adopt or resist algorithmic input (Kesavan and Kushwaha 2020, Balakrishnan et al. 2022, Cui et al. 2022). Central to this stream is algorithm aversion, which characterizes how human trust fractures upon observing algorithmic error (Dietvorst et al. 2015, 2018). Although decision-makers willingly rely on algorithmic forecasts in the absence of performance feedback (Logg et al. 2019), Dietvorst et al. (2015) show that once individuals observe a superior algorithm commit even a single minor error, they abandon it in favor of a human forecaster who makes larger, more frequent mistakes. When the algorithm shifts from an advisory decision aid to an active bargaining counterpart, however, human resistance becomes more nuanced. For instance, Chen and Huang (2026) find that bargaining outcomes against AI counterparts align closely with the full-information, human-to-human behavioral benchmarks established by Davis and Hyndman (2019). Furthermore, **?** find no evidence of intrinsic human favoritism in counterpart selection, showing instead that realized profitability differences in prior rounds govern subsequent counterpart preferences. Our setting differs from classic algorithm aversion, as principals do not have the chance to observe the Bot making possible errors before choosing. We therefore use the broader term AI delegation avoidance to describe the reduced likelihood of delegating to a bot, without attributing this behavior specifically to observed algorithmic failure.

Nevertheless, documenting how an AI negotiator behaves and how a human counterpart reacts leaves a fundamental operational question unresolved: how managers choose to deploy AI when delegating tasks where a machine may not generate the same relational value as a human. Prior research examines Human-Bot delegation almost exclusively within structured, objective task domains such

as image classification and performance forecasting (Fügener et al. 2022, Bockstedt and Buckman 2025, Fügener et al. 2025). In these quantifiable settings, Fügener et al. (2022) find that suboptimal human-to-AI delegation stems from deficient metaknowledge, an inability among humans to accurately assess their own capabilities across task instances, rather than conscious algorithm aversion. Conversely, Bockstedt and Buckman (2025) show that under standard gain framing, principals exhibit algorithm aversion by preferring human experts over demonstrably superior AI assistants. At the same time, reframing task incentives around potential losses heightens situational awareness, mitigating this aversion and driving rational, performance-driven delegation (Bockstedt and Buckman 2025). However, a substantial gap remains regarding how delegation operates in unstructured, interpersonal interactions governed by socio-emotional dynamics. Behavioral evidence indicates that task subjectivity heightens human favoritism because decision-makers believe algorithms lack the intuitive, emotional, and affective capabilities required for interpretive domains (Castelo et al. 2019). In business-to-consumer (B2C) service and commercial interactions, consumers explicitly prefer human agents over AI when exchanges involve subjective rather than objective information (Dai et al. 2026). Similarly, in conversational commerce, counterparts penalize disclosed chatbots because they perceive them as less empathetic, devoid of personal feelings, and less trustworthy (Luo et al. 2019), while the diminished social presence of chatbot interactions reduces willingness to re-engage in future exchanges (Lei et al. 2021). Unlike consumers who directly experience an interaction, in business-to-business (B2B) negotiations principals delegate the interaction but retain its economic consequences. Their choice may therefore reflect not only expected agent competence but also beliefs about how the agent will represent the firm and affect the counterpart's willingness to engage relationally. Because B2B negotiations are inherently unstructured exchanges evaluated across both instrumental and socio-emotional dimensions (Curhan et al. 2006, Vaccaro et al. 2025), principals must trade off the efficiency gains of an AI negotiator (van Hoek and Lacity 2023) against the relational value a human negotiator may generate.

2.2. General- or Relationship- Investments in the Supply-Chain

Delegating a B2B negotiation is not only a choice of who bargains on the firm's behalf, it is also an investment decision that shapes the buyer-supplier relationship. Assigning a dedicated human negotiator commits costly resources to a single counterpart, whereas a negotiation bot is a transferable capability that can be redeployed across many suppliers. The supply chain literature on precontractual investments offers a natural lens for this choice. It distinguishes investments that strengthen a firm's bargaining position in the market from investments that help build relationships

(Beer et al. 2018). This subsection draws on that literature in three steps. We first position AI delegation as a general investment and human delegation as a relationship-specific investment. We then review how relationship-specific investments create value. Finally, we derive our main hypothesis and exploratory questions.

In the supply chain context, deciding whether to deploy an AI negotiator or to delegate to a human agent entails a fundamental economic trade-off between guaranteed process efficiency and uncertain relational returns, which we formalize by adapting the distinction between general and relationship-specific precontractual investments developed by Beer et al. (2018) to the context of AI delegation. Under this framework, precontractual investments fall into two distinct categories: general investments, which represent transferable capabilities that strengthen a firm's outside options and baseline bargaining power, and relationship-specific investments, which represent costly commitments tailored exclusively to a focal partner (Dyer and Singh 1998, Beer et al. 2018). We posit that deploying a negotiation bot functions as a modern general investment. Field evidence from Alibaba's supplier base demonstrates that procurement requests bearing a "smart" automation label secure lower price quotes than RFQs carrying either a human label or an unlabeled automation tag (Cui et al. 2022). Suppliers attributed these concessions to the belief that AI sophistication signals strong competitive outside options, assuming the algorithm possesses the analytical capacity to evaluate and solicit hundreds of alternative vendors simultaneously. Viewed through the theoretical lens of Beer et al. (2018), automating supplier interactions constitutes a general investment because it expands the buyer's outside options. In contrast, recruiting and dedicating human resources to manage a specific counterpart constitutes a relationship-specific investment (Beer et al. 2018). Such dedicated investments carry a direct opportunity cost and may yield positive returns when the counterpart reciprocates.

A costly relationship-specific investment, such as dedicating human personnel to serve a corporate partner, may communicate commitment and can function as a credible signal of trustworthiness (Beer et al. 2018). Because relationship-specific investments commit resources that hold unique value solely within a focal partnership (Dyer and Singh 1998), they serve as costly signals of trust that induce suppliers to reciprocate through higher effort and the provision of non-contractible quality (Beer et al. 2018). Similar to the effort-related findings of Fehr et al. (1993) in the labor market. In turn, these relational commitments directly reduce transaction costs by replacing costly formal contracting and oversight with informal, self-enforcing safeguards, such as goodwill trust (Dyer and Singh 1998, Baker et al. 2002).

In supply chain exchanges, buyers typically evaluate product quality across two distinct governance regimes: ex-ante standardization and ex-post verification. On the one hand, ex-ante standardization relies on general factory-floor certifications (e.g., ISO 9000 (Beer et al. 2018)). While such standardization may be reflected in the unit costs, it establishes contractible baseline compliance, allowing market prices to convey relevant information and keeping ex-post transaction friction low (Dyer and Singh 1998). On the other hand, when production is uncertified or low-cost, product quality remains non-contractible, only observable to the exchange partners ex post (Baker et al. 2002). Under this regime, ensuring quality compliance without relational governance substantially inflates ex-post transaction costs as it requires extensive ex-post auditing, batch testing, and formal legal safeguards (Dyer and Singh 1998, Beer et al. 2018). Establishing a collaborative relationship resolves this verification friction. A trustworthy supplier reciprocates relational investments by delivering higher non-contractible quality (Beer et al. 2018), meaning that both parties can lower transaction costs by relying on relational governance rather than costly monitoring (Dyer and Singh 1998). We operationalize this tension in our experimental design by mapping the relationship-specific investment to Agent A (human) and the general investment to Agent B (human or AI). When the general-investment option (Agent B) is an artificial bot rather than a human, we expect principals to shift behaviorally toward the relationship-specific human alternative.

HYPOTHESIS 1. *If a human principal has the choice between delegating to human Agent A (relationship) and Agent B (general), the likelihood of choosing Agent B is smaller if that Agent B is an AI rather than a human.*

Beyond testing this baseline delegation effect, we conduct exploratory analyses across three operational dimensions to identify when AI delegation avoidance manifests and how it shapes firm profitability. First, because precontractual investments operate as sequential signals of trust (Beer et al. 2018), we examine whether delegation behavior differs between first movers, who initiate the Stackelberg game under strategic uncertainty, and second movers, who choose after observing their counterpart's delegation decision. Second, because optimal deployment in practice depends on the underlying economic environment, we evaluate delegation across Low-, Mid-, and High-Transaction Cost (TC) product regimes (Dyer and Singh 1998, Beer et al. 2018). Whereas a general-investment signal is economically advantageous in Low-TC settings, where formal ex ante certification entails higher unit production costs but minimal ex post transaction costs. A relationship-specific signal dominates in High-TC settings, where low-cost, uncertified production necessitates substantial ex post verification and monitoring expenses. Comparing delegation

across these regimes reveals whether principals avoid AI unconditionally or tailor their avoidance to environments where economic fundamentals favor relationship building. Third, we investigate the downstream profit implications of these delegation choices by comparing realized bargaining outcomes across human-to-human, human-to-bot, and bot-to-bot dyads. Specifically, whereas human-to-human negotiations persistently deviate from the Nash bargaining equilibrium by ordering inefficiently low quantities (Davis and Hyndman 2019), a behavioral inefficiency that persists in human-to-bot interactions (Chen and Huang 2026), we examine whether alternative bot configurations mitigate this coordination failure. Unlike Chen and Huang (2026), who calibrate their bot using human-derived reservation profits, we anchor the bot's reservation profit directly at the theoretical Nash bargaining equilibrium (3.4). Accordingly, we explore whether bot participation improves channel coordination sufficiently to offset, or even surpass, the transaction-cost savings generated by human-to-human relational governance.

3. Methodology

In this section, we first describe the supply-chain context, then we go in-depth into the first stage of the experiment where managers delegate an upcoming negotiation to one of two agents. subsequently, we cover the the second stage which involves the agents negotiating. Finally, we describe the experimental protocol.

3.1. The Supply-Chain Context

We conducted a two-stage lab experiment in which participants were assigned to be employed as manager or a negotiation agent by a retailer or supplier. The two companies are negotiating on the wholesale price w and quantity q of a product. The setting is related to the two-tier supply chain from Davis and Hyndman (2019). A supplier produced q units for production cost $c = \{3, 4, 5\}$, sells them to a retailer for wholesale price w who, in turn, sells D units to final customers for a fixed retail price $r = 11$. There is full information regarding both r and the realization of c, yet the demand was uniformly distributed on $[0, 100]$. In many e-commerce and drop-shipping arrangements, suppliers and retailers negotiate wholesale contracts where the supplier retains ownership of the stock and fulfills customer orders directly. Consequently, while both firms remain exposed to uncertain consumer demand, the supplier unilaterally bears the financial risk of excess inventory whenever the stocked quantity exceeds realized demand ($q > D$).

In addition to the unit production cost $c \in \{3, 4, 5\}$, we introduce a transaction cost $t \in \{50, 40, 30\}$ that is inversely paired with c, yielding three distinct cost profiles: $(c = 3, t = 50)$, $(c = 4, t = 40)$,

and $(c = 5, t = 30)$. Because the downstream retail price remains fixed across all profiles ($r = 11$), this inverse pairing creates a direct tension between the unit margin ($r - c \in \{8, 7, 6\}$) and the fixed transaction overhead t. Practically, this parameterization captures procurement settings with varying verification requirements, ranging from high ($(c = 3, t = 50)$) and mid ($(c = 4, t = 40)$) to low ($(c = 5, t = 30)$) Transaction Cost, where lower unit production costs require higher administrative, auditing, or safeguarding expenses to execute the contract, whereas higher-cost, standardized production reduces downstream contracting friction (Dyer and Singh 1998, Beer et al. 2018). However, while this operational trade-off informed our parameter calibration, we presented c and t to participants strictly as factual cost parameters without any underlying quality or verification narrative. We adopted this neutral framing deliberately to prevent unmodeled behavioral confounds, such as subjective perceptions of product defect risk or quality shirking.

3.2. Stage 1: The Delegation Decision

Like a Stackelberg game, the first mover, without knowing the counterpart's choice, decides to make a relationship-specific investment by delegating to agent type-A, or if they delegate to agent type-B they make a general investment. Second movers, observe the first-mover's choice, and then make their own delegation decision. Type-A agent carries higher delegation cost ($d_A = 20$) compared to type-B ($d_B = 10$). Additionally, if a deal is made, the interaction induces transaction cost t. These transaction cost are halved in case a relationship is built. A relationship can be created if both managers delegate to type-A agents. In that case, there is a 50% chance of creating the relationship and reducing transaction cost to $t/2$ (i.e., trust is created, reducing the needed control mechanisms). This reduction in transaction cost determines the expected cost gap between delegating to agent A and B. For High- (Low-) Transaction Cost making the relationship-specific investment (type-A agent) yields lower (higher) expected cost than the general investment (type-B agent). The mid-Transaction Cost class yields equivalent expected cost.

In our control condition, managers delegate to either human agent type-A or human agent type-B. In the treatment condition, type-B agent is an Bot agent while type-A agents retains it human nature. Managers were financially incentivized to make the decision that maximizes their company profits and they may receive a dividend of 2.5% from the company profits. The stylized company profits were equal to the profits from deal after demand is known minus the negotiation agent bonus and delegation costs. The delegation costs include the base agent cost ($b_A = 20$; $b_B = 10$) and the transaction cost $t \in \{50, 40, 30\}$. In both conditions, the negotiation agent earned a bonus equal to 2% of the realized deal profit. When a manager delegated to the Bot agent, this 2% bonus

was not lost, instead, it was paid to a human participant assigned to the role of an "Bot agent deployer" who monitored the system's performance. This design ensures that a human participant receives the deal bonus, ruling out confounding social preferences in the manager's delegation choice. Having a human receiving the earnings of an artificial agent has been proved to be essential across bargaining and economic dilemma games to avoid the confounding effects of ambiguous recipient of the machine earnings (von Schenk et al. 2023, Fu et al. 2026).

3.3. Stage 2: The Negotiation

After retailer and supplier managers decided to delegate to type-A or type-B agents, actual negotiations took place. Depending on the treatment and the delegation decisions these could be human-human, human-bot or bot-bot negotiations.

The negotiation interface consisted of three features: a chatbox for natural language instant messaging exchange, two numerical boxes for binding offers only, and a decision support tool. While the decision support tool provides negotiators with the price and quantity of the Nash Bargaining solution, they can input other terms to check expected profit for their counterpart and for themselves (see EC for snapshot and reusable code). The participants have five minutes to reach an agreement. Each party can send a message or make a binding offer at any time. Any new binding offer pulls the old (own) offer off the table. Both parties can accept the counterpart's most recent binding offer any time by clicking accept. If no agreement is reached within the allotted time, both parties receive zero profit for that negotiation. This interface is consistent between treatments, meaning humans that negotiate with bot counterparts see the same interface than humans that negotiate with other human counterparts.

3.3.1. The Bot Counterpart The primary objective of the bot is to secure a reservation profit determined by the Nash bargaining solution. This being the Pareto-optimal symmetric combination of wholesale price (w) and order quantity (q) that maximizes joint expected channel profit while minimizing expected profit differences between the two parties. While the Bot agent is programmed to propose complete Pareto-optimal offers (quoting both q and w), it also possesses the flexibility to process, adapt to, and earn higher profits from partial (single-term) offers made by the counterpart. Extending the work of Chen and Huang (2026), which restricted bargaining to complete offers, our agent evaluates whether a counterpart's single-term proposal can yield an expected profit at least equal to the reservation profit. If a partial offer can achieve or exceed this benchmark, the agent adapts its counter-offer accordingly; otherwise, it defaults back to proposing the Nash bargaining

solution (see E-Companion for detailed calculations). Note that this interactive bargaining logic is only required in bot-human negotiations, as bot-bot negotiations automatically settle on the Nash bargaining equilibrium.

To execute this strategy in unstructured chat, we adapted the hybrid (rule-based and large language model) architecture of Garrido Tenorio et al. (2026). Because our full-information setting allows for open-ended bargaining without pre-scripted rule-based messages, the only hard-coded rule we retain and refine is an automated guardrail check prior to message transmission. Specifically, the system extracts the proposed terms from the Bot agent's tentative message and calculates the implied expected profit; if this profit falls below the pre-defined reservation profit, a regeneration loop forces the model to produce a new compliant message. Technically, the Bot agent performs three natural language tasks: two natural language processing (NLP) tasks to parse incoming and drafted messages into formatted numerical values, and one natural language generation task to compose instant chat replies. The message generator relies on a dynamic system prompting technique (Garrido Tenorio et al. 2026) comprising three static components: (1) the assigned role (retailer or supplier), (2) the relevant unit cost (c) or retail price (r), and (3) the negotiation rules (see E-Companion). At each turn of the conversation, the model is dynamically fed: (1) the ongoing chat transcript via user prompts, (2) the calculated optimal offer, (3) few-shot examples of ideal responses, and (4) instructions to adhere to the negotiation rules and conversation history.

3.4. Experimental Protocol

We ran the experiments over two months in 2026 at the behavioral lab of a large European public university. The sample consisted of 500 undergraduate and graduate students with ages ranging between 18 and 25. Participants received a show-up fee of 5€ and were additionally financially compensated based on their performance. The total payment ranged between €5 and €18,7. The average duration of the experiment was 18 minutes. Participants received instructions both on paper and on screen. After reading those, they had to complete 5 comprehension checks: negotiators calculated the expected cost of given deals, and managers calculated (expected and actual) delegation and transaction costs under different set-ups (See E-Companion). First, we assign the whole cohort of four participants into the control or treatment condition, second we randomly assign roles to participants, two managers and two negotiators, third we randomly assign one of three product classes (High-, Mid-, Low-Transaction Cost) to the cohort. Lastly, we randomly assign retailer and supplier managers to be the first or the second mover of the stackelberg game. After the negotiation was conducted all participants received their final compensation. We collected the comprehension

attempts, delegation decision, conversation, offers and final negotiation results. The experiment was implemented through the experimental software oTree (Chen et al. 2016). For replication purposes an open-source github repository is available in the E-Companion.

In total 500 subjects participated in the experiment. 250 as managers and 250 as negotiators. Participants with the role of managers who failed the comprehension check more than three times (n=8) were excluded from the analysis. A total of 242 managers are included in our main analysis. As post-hoc, we analyze deal performance, since 5 participants with the role of negotiators failed the comprehension check more than three times we also exclude them from the post-hoc. As robustness check, we also conducted all the analyses without any exclusion, yielding qualitatively similar results as reported in the E-Companion. We pre-registered the hypotheses, experimental design, analyses, and exclusions at the Open Science Framework (OSF).

4. Results

In this section we first test our Hypothesis, and subsequently explore the first/second mover decision dynamics. Finally, we dive into the heterogeneous treatment effects by product class.

4.1. Main Effect of AI Treatment

We test the main effect of the AI treatment using logistic regression analysis with a binary dependent variable: the choice to delegate to Agent B (general investment) versus Agent A (relationship investment). Given our directional prediction that AI presence in the general investment reduces the likelihood of choosing Agent B, we report one-sided p-values for the AI Treatment coefficient in our primary test, consistent with our pre-registration. Our sample consists of the 242 subjects who passed the comprehension checks.

As presented in Table 1, the positive AI treatment coefficient ($\beta = 0.383$, $SE = 0.265$) is weakly significant in the hypothesized direction ($p = 0.073$, one-sided). This corresponds to an odds ratio of 1.47, indicating that the odds of choosing the relationship-investment option are approximately 47% higher in the AI-treatment condition than in the human-only condition. Model 2 controls for the experimental manipulation that varies the reward-risk tradeoff between relationship-specific and general investment. The estimated treatment effect remains similar after controlling for product class ($\beta = 0.390$, $SE = 0.270$, $p = 0.074$, one-sided), providing initial support for Hypothesis 1.

Because delegation decisions were made sequentially, as a Stackelberg game, Models 3 and 4 examine the AI-treatment effect separately for first and second movers. Model 3 restricts the sample to first movers who passed the comprehension checks ($N = 119$). To gain a better understanding of

Table 1 Logistic Regression

	DV: Relationship Investment (RI)			
	(1)	(2)	(3)	(4)
Bot as General Investment	0.383*	0.390*	0.264	0.621*
	(0.265)	(0.270)	(0.388)	(0.478)
Low Transaction Cost Product Class		−0.506	−0.478	−0.166
		(0.320)	(0.468)	(0.557)
High Transaction Cost Product Class		0.518	0.104	1.413**
		(0.339)	(0.484)	(0.629)
First Mover Chose RI				3.028***
				(0.531)
Constant	0.249	0.260	0.609	−2.290***
	(0.183)	(0.260)	(0.381)	(0.630)
Observations	242	242	119	123
Log Likelihood	−161	−156	− 76	−56
AIC	326	321	160	122

Note: *p<0.1; **p<0.05; ***p<0.01
Treatment's P-value is halved due to the pre-registered directional hypothesis. Model 2 controls for the product class. Model 3 is subsets on first mover. Model 4 subsets on second mover and controls for first mover.

how the second mover reacts to the counterpart's initial decision, Model 4 in Table 1 restricts the sample exclusively to second movers, controlling for the first mover's choice. While this roughly halves the sample size ($N = 123$), it is important to note that we retain slightly more second movers than first movers because a higher proportion of first movers failed the comprehension check. As the first mover's observed decision functions merely as a fixed binary input for the second mover, we do not exclude valid second-mover observations simply because their counterpart in the dyad failed the comprehension check.

Figure 1 shows the descriptive statistics of the delegation decision of first and second movers between treatments. The distribution of second-mover responses reveals a distinct shift in reciprocity patterns contingent upon the AI general investment treatment. When the first mover initiates a relational (uncertain) investment, second movers exhibit a higher frequency of reciprocal relational choices under the AI treatment (82.9%) compared to the human control condition (75.7%). Conversely, when the first mover pursues a general (certain) investment, the rate of reciprocal general choices is lower in the AI treatment (76.2%) than in the human control (87.5%).

Among first movers, the treatment coefficient is positive but not statistically significant ($\beta = 0.264$, $SE = 0.388$, $p = 0.248$, one-sided), providing no clear evidence of an AI-treatment effect. However,

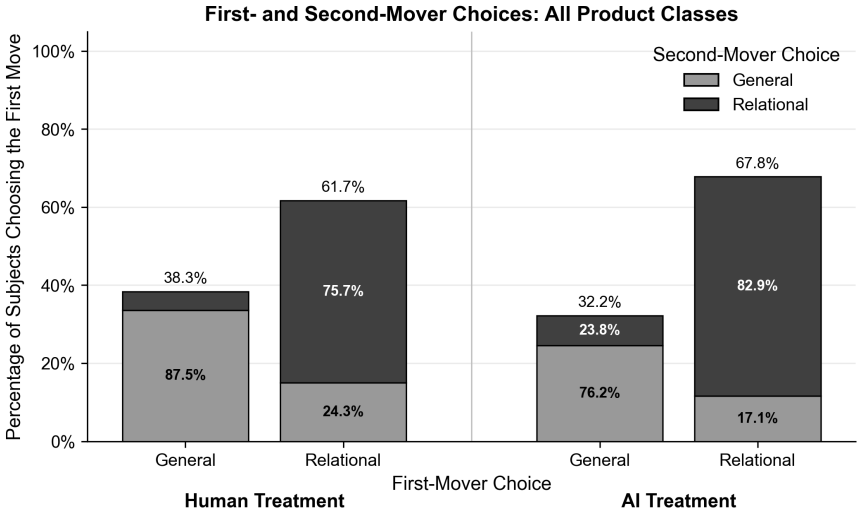


Figure 1 Participants Choosing Relationship Investment by Treatment and Product Class

an analysis of average marginal predictions based on Model 3 reveals that the overall probability of choosing the relationship investment (Agent A) is 64.9% (95% CI [55.9%, 73.0%]). Because this interval is strictly above 50%, it suggests that first movers generally exhibit a baseline tendency toward relationship investments, independent of the treatment condition.

Among second movers (Model 4), the estimated treatment coefficient is larger but remains only weakly significant under the preregistered one-sided test ($\beta = 0.621$, $SE = 0.478$, $p = 0.097$, one-sided). Holding product class and the first mover's choice constant, the AI treatment is associated with approximately 86% higher odds of choosing the relationship-investment option. To test whether this effect was driven specifically by second movers reciprocating a relationship investment, we ran an exploratory model including an interaction term between the AI treatment and the first mover's choice. The interaction was not statistically significant ($p = 0.621$, see Appendix), indicating that the AI treatment increases the likelihood of relationship investment among second movers generally, regardless of the first mover's initial action.

Overall, the coefficients are consistently positive across all four specifications, as predicted. However, the evidence is modest: the treatment effect is weakly significant in the full sample and among second movers under the preregistered one-sided tests, but it is not statistically significant among first movers. Generally, the results are in line with Hypothesis 1.

4.2. Effect of AI Treatment within Product Classes

Since not every product class rewards relationship-building investments equally, we conduct a manipulation check to assess if the assigned product class affected the risk-taking behavior of participants. Figure 2 visually shows how the proportion of participants making the risky relationship

Table 2 Logistic Regression by Product Class

	DV: Relationship Investment (RI)		
	Transaction Cost levels		
	High	Mid	Low
	(1)	(2)	(3)
Bot as General Investment	0.837*	0.331	0.100
	(0.516)	(0.455)	(0.447)
Constant	0.580*	0.288	−0.100
	(0.334)	(0.312)	(0.317)
Observations	80	82	80
Log Likelihood	−46	−55	−55
Akaike Inf. Crit.	95	113	115

Note: *p<0.1; **p<0.05; ***p<0.01 Treatment's P-value is halved due to the pre-registered directional hypothesis.

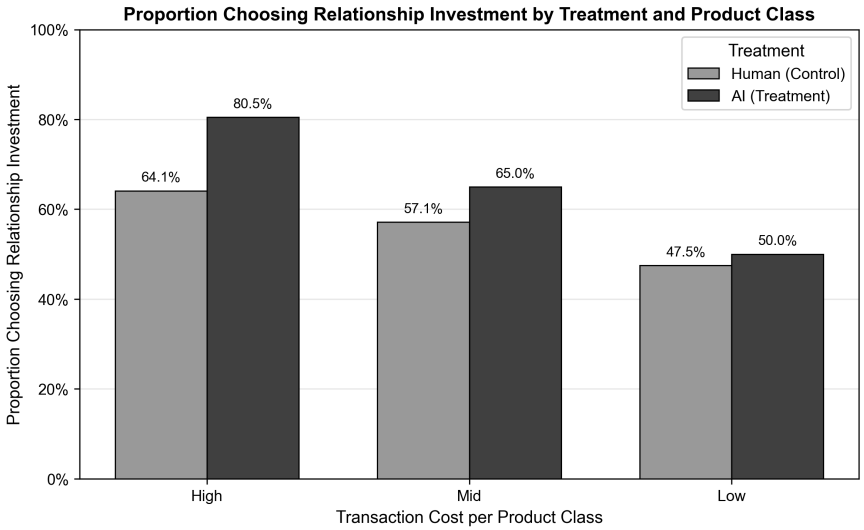


Figure 2 Participants Choosing Relationship Investment by Treatment and Product Class

investment is highest in the High-Transaction Cost product class and lowest in the class with Low-Transaction Cost.

Using chi-squared tests, we find a significant difference in the proportion of participants choosing the relationship investment between the High-TC and the Low-TC class ($\chi^2(1, N = 160) = 8.48$, p = .004). However, the difference between those classes with the Neutral class is not significantly different.

To explore whether the AI treatment effect and the risk-taking behavior varied depending on the economic context of the negotiation, we subset our sample and ran separate logistic regressions for the High-, Mid-, and Low-TC product classes (Table 2).

First, an analysis of average marginal predictions reveals that managers rationally adjusted their risk-taking behavior in response to the assigned transaction costs. The overall probability of choosing the relationship investment (Agent A) was highest in the High-TC class (73.1%, 95% CI [62.1%, 81.8%]), where the expected cost favored building a relationship. In the Neutral class, where the expected cost between options is strictly equal, managers still exhibited a baseline preference for relationship-investment (61.1%, 95% CI [50.2%, 71.1%]). However, in the Low-TC class, where the expected cost penalized the relationship investment, this risk-taking tendency vanished, dropping to chance levels (48.7%, 95% CI [38.0%, 59.6%]).

Against these rational baselines, we find that the magnitude of the AI treatment effect scales directly with the economic alignment of the product class. While the presence of an AI counterpart in the general investment option consistently increased the probability of choosing the relationship investment across all conditions, this effect was only pronounced in the High-TC class. Figure 2 shows that relational investment offers a superior though uncertain expected return, the AI treatment increased the probability of choosing Agent A by 16.4 percentage points (from 64.1% to 80.5%), a weakly significant increase under our preregistered one-sided test ($\beta = 0.837$, $SE = 0.516$, $p = 0.053$, one-sided).

While the models in Table 2 on the Product Class subset do not reveal statistically significant heterogeneity in the AI treatment within most product classes, the treatment effect within the High-TC class is weakly significant ($p = 0.052$, one-sided). This exploratory post-hoc analysis indicates that AI aversion is highly contingent on the economic stakes of the transaction. Subjects demonstrate a willingness to forfeit cost certainty to avoid delegating an interaction with counterpart to AI only when the relational alternative provides a superior expected return (the High-TC class). However, when the general investment offers equal (Neutral class) or superior (Low-TC class) expected cost, the AI treatment effect is statistically indistinguishable from zero ($p = 0.23$ and $p = 0.41$, respectively one-sided). Ultimately, while participants exhibit some AI aversion, they are unwilling to sacrifice equal or strictly higher certain profits merely to avoid delegating an interaction with counterpart to AI.

4.3. Post-hoc Treatment Effects on Company Profit

The negotiation is held right after the delegation decision. This section analyses the negotiation outcomes, which are a consequence of the managerial choice to invest in a potential relationship or in a general investment. The actual negotiation outcomes are subject to random demand $D \sim U[0, 100]$, and if both sides choose to invest in the relationship there is a coin flip to determine

if the relationship is realized. This post-hoc analysis leverages the actual delegation decision and assumes the expected values of transaction cost, we also use the actual deal terms to calculate the expected deal profits.

Our experimental manipulation switches the human agent in the General Investment for a bot that negotiates with the counterpart on behalf of the delegator. Thus all, negotiations in the control conditions are between human agents. In the AI treatment, of the 125 negotiations, 62 occurred in Human Control and 63 in the AI Treatment. Within the AI Treatment, 35/63 (55.56%) paired two human agents and 16/63 (25.40%) paired two AI agents. The remaining negotiations were mixed: 7/63 (11.11%) involved a human retailer and an AI supplier, while 5/63 (7.94%) involved an AI retailer and a human supplier. Table EC.1 reports agreement rates by agent pairing and product class, pooling Human–Human negotiations across conditions. Two-sided Fisher exact tests do not detect an agreement-rate difference between Bot-Bot and Human-Human negotiations ($p = 0.209$), between the two mixed pairings pooled and Human-Human negotiations ($p = 0.67$), or between any negotiation with at least one bot and a Human-Human negotiation ($p = 0.518$). The Bot-Bot group reached agreement, as programmed in all 16 cases, but several comparison cells are small and the observed pairings were chosen by managers.

Table 3 reports accepted prices and quantities and the corresponding expected retailer, supplier and joint deal profits under $D \sim U[0, 100]$. We calculate expected profit at each accepted deal's own price and quantity and then average within each agent pairing and product class. The joint deal profit is the sum of the two firms' deal profits before delegee bonuses, delegation costs and transaction costs. Conditional means use agreements only; the unconditional joint mean assigns zero deal profit to failed negotiations.

Among agreements, each observed class-specific dyad with at least one bot has a higher mean expected joint deal profit than the Human-Human dyad in the same class. These are descriptive comparisons: the mixed-dyad cells are small, and managers selected the delegation arrangements. The cost trade-off is clearest in class A. When both managers choose the Relationship Investment, the relationship lottery saves the pair 25 in expected transaction costs relative to two General Investments, while the two Relationship Investments add 20 in delegation costs. The net cost advantage is therefore 5 per agreement. Among class-A agreements following two Relationship choices, the 23 Human-Human negotiations yield a mean expected joint deal profit of 263.94; the three Bot-Bot negotiations following two General choices reach the Nash terms and yield 290.91. The observed deal-profit gap of 26.97 exceeds even the gross expected transaction-cost saving of

Table 3 Expected deal profit and company-profit components by pooled agent pairing and product class.

		Terms		Deal profit by role		Joint deal profit		Costs			Joint company profit	
Agent pairing	Agreements/N	Price	Quantity	Retailer	Supplier	All	Agreement	Transaction	Base	Bonus	All	Agreement
High Transaction Cost												
Human–Human	30/36	6.93	56.43	155.70	95.76	209.55	251.46	67.36	36.11	4.19	101.89	128.93
human–bot	2/2	7.70	66.50	145.65	140.81	286.45	286.45	100.00	30.00	5.73	150.72	150.72
Bot–Bot	3/3	7.86	73.00	145.55	145.35	290.91	290.91	100.00	20.00	5.82	165.09	165.09
Mid Transaction Cost												
Human–Human	28/32	7.17	52.43	140.76	56.16	172.31	196.92	58.13	34.38	3.45	76.35	91.56
human–bot	4/6	8.01	54.75	112.26	96.07	138.89	208.34	53.33	30.00	2.78	52.78	94.17
Bot–Bot	4/4	8.43	64.00	111.85	110.87	222.72	222.72	80.00	20.00	4.45	118.27	118.27
Low Transaction Cost												
Human–Human	26/29	8.00	53.35	106.77	28.41	121.20	135.18	46.55	32.41	2.42	39.82	48.25
human–bot	4/4	9.37	63.25	65.77	81.78	147.55	147.55	60.00	30.00	2.95	54.60	54.60
Bot–Bot	9/9	8.94	55.00	82.14	81.48	163.63	163.63	60.00	20.00	3.27	80.36	80.36

Note: The unit is the negotiation. Terms, deal profits by role, and columns marked Agreement average agreements only. Columns marked All and the three cost columns average every negotiation. After failure, deal profit, transaction cost, and agent bonus are zero; base delegation cost remains payable. The bonus is the financial incentive and compensation to the human agent, one-fiftieth of the firm's deal profit. Joint company profit equals joint deal profit minus joint transaction cost, joint base cost, and bonuses. The relationship lottery gives expected transaction cost $0.75\,TC_k$ per firm after a Relationship–Relationship agreement and TC_k otherwise. Displayed All company profit is calculated from the displayed two-decimal components so each row adds exactly. Pairings are the managerial delegations. human-bot pools both mixed role assignments within each class.

25. Thus, in these observed class-A negotiations, the expected relationship saving does not offset the lower value of the deals reached by human agents.

The conclusion of this section is to say that even if managers make the optimal decision by making the relationship specific investment, because the human delegee when it negotiates with anotehr human reaches deals that in expected profit are lower than the deals reached when an AI is involved in the negotiation. Specially since this difference is larger than the expected gain from the saving in transaction cost.

4.3.1. Impact of AI Agents Negotiator on Deal Terms We pool Human-Human negotiations across Human Control and the AI Treatment, combine the two directions of mixed Human-Bot negotiation, and report Bot-Bot negotiations separately. Table 4 uses accepted agreements only. For each agreement, the signed quantity and price gaps equal the accepted term minus its product class's Nash term; absolute gaps measure distance from Nash. For the profit comparison below, expected joint deal profit is calculated from each accepted quantity under $D \sim U[0, 100]$ before averaging.

Human-Human agreements order 10.29 units below the class-specific Nash quantity on average, compared with 1.70 units below Nash for mixed Human-Bot agreements. Their mean absolute quantity gaps are 19.10 and 10.50 units, respectively (two-sided Welch $p = 0.057$). Mixed agreements are also closer in wholesale price: their mean absolute price gap is 0.43, compared with 1.31 for Human-Human agreements ($p = 0.001$). All 16 Bot–Bot agreements coincide with the Nash price and quantity by design.

Table 4 Deviations from Class-Specific Nash Terms by Agent Pairing

	Agent pairing				
	Human–Human	Human–Bot	Bot–Bot	Mixed − Human	p-value
Signed quantity gap	−10.29	−1.70	0.00	8.59	0.152
Absolute quantity gap	19.10	10.50	0.00	−8.60*	0.057
Signed price gap	−1.04	−0.03	0.00	1.01***	0.002
Absolute price gap	1.31	0.43	0.00	−0.88***	0.001
Agreements	84	10	16		

Note: *p<0.1; **p<0.05; ***p<0.01. Human–Human pools both experimental conditions; Human–Bot pools both mixed role assignments. Signed gaps are accepted terms minus class-specific Nash terms; absolute gaps measure distance from Nash. Differences and two-sided Welch p-values compare mixed with Human–Human agreements. Bot–Bot terms equal Nash by design.

The raw mean expected joint deal profit is 197.29 for Human-Human and 199.65 for mixed Human-Bot agreements, a small difference that is not statistically significant (Welch $p = 0.907$). The groups' product-class composition differs, however, as shown by their mean class-specific Nash deal profits of 228.78 and 212.72. Relative to that benchmark, Human-Human agreements lose 31.49 in expected joint deal profit on average, compared with 13.07 for mixed agreements ($p = 0.040$); Bot-Bot agreements have zero loss because they implement Nash. The deal-profit table (Table EC.2) also shows higher mean expected joint profit for mixed than for Human–Human agreements within each observed product class, although the mixed class-A cell has only two agreements.

The smaller Nash loss is consistent with mixed negotiations choosing quantities closer to the coordinated benchmark. Wholesale price affects how deal profit is divided between retailer and supplier but cancels from their joint deal profit at a given quantity. Ordering below Nash may reflect concern about unsold inventory, yet the data do not identify negotiators' motives. These comparisons condition on agreement and on managers' chosen delegation arrangements; they describe the realised groups rather than a causal effect of assigning an AI negotiator.

5. Discussion

We investigated whether the artificial nature of a general investment behaviorally nudges managers toward relationship-specific investments in B2B negotiations. We find that managers' decision to invest in relationship-specific when the general investment comes from an AI depends systematically on the risk-reward tradeoff between the general and relationship-specific investment, and on the first mover's relationship-building signal. When AI represents a safe, general investment, managers become more willing to embrace the risks of a human relationship-specific investment. This is significant in the high-TC product class, where human relationships yield higher expected returns, and is not significant (yet directionally consistent) in the low-TC category, where the general investment is objectively the superior choice. Finally, when the first mover signals the willingness to build a relationship, the second mover choice to reciprocate the risk taken by the first mover amplifies significantly when AI represents the general investment.

Our finding that the mere presence of AI in the general investment option pushes managers away from it toward the riskier, human relationship-specific alternative speaks to a growing body of work on AI adoption in B2B negotiations. Building on prior field evidence that automated RFQs signal stronger outside options to counterparts eliciting steeper discounts than human agents (Cui et al. 2022). We find that managers are roughly 32% less likely to choose the general investment when it entails delegating a negotiation to a bot rather than a human agent. This effect persists after controlling for the product class manipulation, and is even more pronounced among second movers. If the first mover initiates with a relationship-specific investment, the second mover's willingness to reciprocate is significantly higher when the alternative general investment is artificial. This aligns with prior literature suggesting that task subjectivity increases algorithm aversion (Castelo et al. 2019), prompting managers to bypass the automated general investment when they feel the context requires a human touch. While AI-driven automation has proven effective for managing tail-end, low-stakes spend (van Hoek and Lacity 2023), our results show that its presence can also pull managers toward, rather than away from, costly relationship-specific investments that function as signals of trust (Beer et al. 2018).

Our experimental design exposed subjects to a high-Transaction Cost (low-TC) scenario in which a risky relationship-specific investment either does (does not) yield a higher expected payoff relative to a certain general investment. We operationalized the relationship-specific investment as a dedicated and costly human resource assigned to manage the negotiation. In our treatment condition, the general investment is an artificial agent negotiating on the manager's behalf; in the

control condition, both the general and relationship-specific investments are human. Operating under a Stackelberg framework, the first mover makes an investment decision, and the fully informed second mover reacts. Following these decisions, the negotiation occurs, the outcome of which determines both parties' final payoffs.

This inclination to bypass AI heavily depends on the expected payoff difference between general and relationship-specific investments, which we operationalize as the product class. When transaction costs are high and human relationships yield higher expected returns, the presence of AI acts as a behavioral catalyst that steers managers toward riskier but potentially more rewarding human partnerships. Here, maintaining human involvement is treated as a strategic necessity, and managers willingly absorb relational risks rather than delegate to a machine. Conversely, in low-tc categories where automated delegation objectively generates higher financial returns, managers do not significantly penalize the artificial agent. When theoretical predictions strongly favor the general investment, algorithmic aversion dissipates, and managers willingly forego risky human connections to capture the guaranteed cost savings and expanded outside options provided by the algorithm.

Ultimately, managers successfully deploy machines to maximize efficiency in product categories with low transaction cost and strategically reserve dedicated human negotiators for critical partnerships where reciprocity and signaling are essential (Beer et al. 2018). A key managerial implication of our work is that granting managers the option to delegate B2B negotiations to AI does not degrade decision quality: managers do not indiscriminately over-delegate to AI when human involvement carries strategic value, nor do they cling to costly human negotiators when AI delegation is objectively the superior choice. In other words, the option to delegate to AI is not a source of systematic bias but rather a tool that managers deploy contingent on the underlying economics of the relationship. Therefore, firms should confidently adopt AI negotiation agents and rely on managerial discretion to delegate B2B interactions.

Yet the relational benefits of the human relationship-specific investment do not compensate for the efficiency loss in human–human deal terms. Although human and artificial negotiators alike had access to the Nash bargaining solution offer, agreements between two human negotiators fall short of the class-specific Nash benchmark by 31.49 in expected joint deal profit, compared with 13.07 when one of the negotiators is a bot ($p = 0.04$), while Bot-Bot agreements implement the Nash terms by design. Humans thus reach more efficient deals when a bot is involved as their counterpart. The loss originates in quantities rather than prices. Human–human agreements order on average 10.29 units

less than the Nash quantity, and their absolute quantity gap is almost twice that of mixed agreements (19.10 versus 10.50 units, $p = 0.057$). The wholesale price only divides the surplus between retailer and supplier, so the lower prices negotiated by human pairs (1.04 below Nash) shift deal profit toward the retailer without reducing total supply chain value. Under-ordering, by contrast, is a conservative response to the supplier's risk of unsold inventory: it leaves demand unserved and shrinks the joint surplus. The relationship-specific investment is too weak a lever to offset this loss. Even when both parties invest in the relationship, the transaction cost is halved only with probability one half, an expected reduction of 25%. Transaction costs amount to roughly one third of the Nash expected joint deal profit (34–37% across product classes), so this reduction lowers the joint expected transaction cost by 25, 20 and 15 in the high-, mid- and low-transaction cost classes, respectively. Net of the 20 in additional base costs that two relationship-specific investments require, the expected joint saving is only 5, 0 and −5. Even when the relationship materialises and the transaction cost is halved, the net joint saving of 30, 20 and 10 remains below the deal profit that human–human agreements forgo relative to Bot-Bot agreements: 39.45, 25.80 and 28.45, conditional on agreement. Accordingly, in the high-TC class, human-human agreements yield a joint expected company profit of 128.93, compared with 150.72 for mixed and 165.09 for Bot-Bot agreements. These comparisons condition on agreement and on the delegation arrangements managers chose, and the mixed High-TC class cell contains only two agreements. They therefore describe realised groups rather than the causal effect of assigning an AI negotiator. Nevertheless, they qualify our behavioral findings. Managers' preference for a human relationship-specific investment in categories with high transaction cost is intelligible as a signal of trust, but in our setting it is the presence of an algorithmic counterpart, not the human relationship, that brings deal terms closer to the efficient supply chain benchmark.

5.1. Strengthts, Limitations, and Future Research

Our experimental design purposefully manipulates the attractiveness of the relationship building option. Each dyad is assigned to one of three product classes: in the product class with high transaction cost, choosing the general investment option yields a higher expected cost than the relationship-specific investment, whereas in the low-TC product class it yields a lower expected cost. The third product class is neutral, with expected costs at par between the two investment options. This setup not only lends supply chain realism, since not all counterparts are strategic, but also allows us to control for risk preferences. Our experimental design therefore enables a clean isolation of the AI treatment.

A primary limitation of our experimental design is the absence of an exit mechanism or a Best Alternative to a Negotiated Agreement (BATNA) for the second mover. In our current framework, when a second mover is approached by an artificial agent, or presented with a general investment, they are compelled to proceed with the negotiation. They do not possess the autonomy to simply reject the automated interaction, opt out in favor of a guaranteed BATNA, or seek an alternative partner who might offer a more desirable relationship-specific investment. In real-world B2B contexts, managers often have the flexibility to walk away from automated negotiations if they strongly prefer human interaction or feel the context demands relational signaling. Future research should incorporate an explicit exit option or alternative partner selection mechanism to better capture how the availability of outside options influences algorithmic aversion and the dynamics of relationship building. A second limitation pertains to the statistical power of our findings regarding the baseline effect of algorithmic aversion. Although the empirical evidence for AI aversion in our sample falls short of conventional statistical significance thresholds ($p < 0.05$), the directional consistency of the negative treatment coefficient across all model specifications, combined with our a priori directional hypothesis, leads us to cautiously interpret these results as supportive of Hypothesis 1. To provide more definitive evidence regarding the presence and magnitude of AI aversion in relational delegation contexts, future research must build upon this foundation by employing larger sample sizes. Enhanced statistical power will be crucial for precisely quantifying the behavioral penalties managers assign to artificial agents and further validating the conditions under which these penalties dissipate. Finally, future research should extend beyond the negotiation phase to explore post-agreement behavioral responses and reciprocal actions. For instance, it remains an open question whether a human counterpart will exert the same level of cooperative effort during contract execution after bargaining with a machine as they would after negotiating with a human. More importantly, subsequent studies should investigate the managerial delegation decision regarding the behavioral design of these automated agents. Specifically, researchers could examine the trade-offs managers face when choosing to delegate to an AI programmed to use deception to maximize surplus versus an AI constrained to remain strictly truthful. While a lying AI might secure highly favorable upfront terms, it exposes the firm to significant risk if the counterpart discovers the deceit. Future work must explore how this discovery triggers retaliation. Analyzing the tension between the immediate financial gains of deploying a deceptive agent and the long-term relational damage caused by such retaliation represents a vital next step in understanding the broader consequences of AI adoption in supply chain partnerships.

References

Baker G, Gibbons R, Murphy KJ (2002) Relational Contracts and the Theory of the Firm. *The Quarterly Journal of Economics* 117(1):39–84, ISSN 0033-5533, URL https://www.jstor.org/stable/2696482.

Balakrishnan M, Ferreira K, Tong J (2022) Improving Human-Algorithm Collaboration: Causes and Mitigation of Over- and Under-Adherence. *SSRN Electronic Journal* ISSN 1556-5068, URL http://dx.doi.org/10.2139/ssrn.4298669.

Beer R, Ahn HS, Leider S (2018) Can Trustworthiness in a Supply Chain Be Signaled? *Management Science* 64(9).

Bockstedt JC, Buckman JR (2025) Humans' Use of AI Assistance: The Effect of Loss Aversion on Willingness to Delegate Decisions. *Management Science* ISSN 0025-1909, 1526-5501, URL http://dx.doi.org/10.1287/mnsc.2024.05585.

Borsani C, Moretto A, Caniato F (2026) Hybrid intelligence in buyer-supplier negotiations: a conceptual framework from literature and market review on AI capabilities. *International Journal of Production Research* 1–31, ISSN 0020-7543, 1366-588X, URL http://dx.doi.org/10.1080/00207543.2026.2626540.

Caro F, Colliard JE, Katok E, Ockenfels A, Stier-Moses N, Tucker C, Wu DJ (2026) Introduction to the Special Issue on the Human-Algorithm Connection. *Management Science* 72(1):1–13, ISSN 0025-1909, URL http://dx.doi.org/10.1287/mnsc.2023.intro.v72.n1.

Castelo N, Bos MW, Lehmann DR (2019) Task-Dependent Algorithm Aversion. *Journal of Marketing Research* 56(5):809–825, ISSN 0022-2437, 1547-7193, URL http://dx.doi.org/10.1177/0022243719851788.

Chen DL, Schonger M, Wickens C (2016) oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* 9:88–97, ISSN 22146350, URL http://dx.doi.org/10.1016/j.jbef.2015.12.001.

Chen Y, Huang R (2026) Haggling with a Bot: Human vs. LLM Negotiation in Supply Chain Contracts, URL https://aspredicted.org/4kb8-kv2d.pdf.

Chen Y, Kirshner SN, Ovchinnikov A, Andiappan M, Jenkin T (2025) A Manager and an AI Walk into a Bar: Does ChatGPT Make Biased Decisions Like We Do? *Manufacturing & Service Operations Management* msom.2023.0279, ISSN 1523-4614, 1526-5498, URL http://dx.doi.org/10.1287/msom.2023.0279.

Cui R, Li M, Zhang S (2022) AI and Procurement. *Manufacturing & Service Operations Management* 24(2):691–706, ISSN 1523-4614, 1526-5498, URL http://dx.doi.org/10.1287/msom.2021.0989.

Curhan JR, Elfenbein HA, Xu H (2006) What do people value when they negotiate? Mapping the domain of subjective value in negotiation. *Journal of Personality and Social Psychology* 91(3):493–512, ISSN 1939-1315, 0022-3514, URL http://dx.doi.org/10.1037/0022-3514.91.3.493.

Dai X, Zhang L, Huang Z, Wang L (2026) AI or human: How the type of information to be disclosed alters customer service agent preferences. *Journal of Retailing and Consumer Services* 89:104621, ISSN 09696989, URL http://dx.doi.org/10.1016/j.jretconser.2025.104621.

Davis AM, Hyndman K (2019) Multidimensional Bargaining and Inventory Risk in Supply Chains: An Experimental Study. *Management Science* 65(3):1286–1304, ISSN 0025-1909, 1526-5501, URL http://dx.doi.org/10.1287/mnsc.2017.2985.

Dietvorst B, Simmons J, Massey C (2015) Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err. *Journal of Experimental Psychology: General* 144(1):114–126, ISSN 0096-3445, URL http://dx.doi.org/10.1037/xge0000033.

Dietvorst BJ, Simmons JP, Massey C (2018) Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. *Management Science* 64(3).

Dyer JH, Singh H (1998) The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management. The Academy of Management Review* 23(4):660–679, ISSN 03637425, URL https://www.proquest.com/docview/210960813/abstract/9583A5710DF74D76PQ/1, num Pages: 20.

Falcão Filho HA (2024) Making sense of negotiation and AI: The blossoming of a new collaboration. *International Journal of Commerce and Contracting* 8(1-2):44–64, ISSN 2976-8829, 2976-8837, URL http://dx.doi.org/10.1177/20555636241269270.

Fehr E, Kirchsteiger G, Riedl A (1993) Does Fairness Prevent Market Clearing? An Experimental Investigation. *The Quarterly Journal of Economics* 108(2):437–459, ISSN 0033-5533, 1531-4650, URL http://dx.doi.org/10.2307/2118338.

Forbes (2026) The AI Agents That Actually Work Don't Post On Social Media. URL https://web.archive.org/web/20260622093417/https://www.forbes.com/sites/guneyyildiz/2026/06/22/the-ai-agents-that-actually-work-dont-post-on-social-media/, section: Energy.

Fu Y, Hanaki N, Wang H (2026) Do Humans Bargain Differently with AI? Evidence from Alternating-Offer Games URL https://aspredicted.org/277540.pdf.

Fügener A, Grahl J, Gupta A, Ketter W (2022) Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research* 33(2).

Fügener A, Schiffels S, Thonemann UW (2025) Project Portfolio Selection with Strategic Buckets—The Role of Naïve Diversification. *Manufacturing & Service Operations Management* msom.2020.0373, ISSN 1523-4614, 1526-5498, URL http://dx.doi.org/10.1287/msom.2020.0373.

Garrido Tenorio K, Becker-Peth M, de Vries J (2026) Hybrid (Rule-Based + Gen AI) Chatbot: Automating Negotiations. Følstad A, Hobert S, Papadopoulos S, Law ELC, Araujo T, Luger E, Brandtzaeg PB, eds., *Chatbots and Human-Centered AI*, 412–427 (Cham: Springer Nature Switzerland), ISBN 978-3-032-26717-7, URL http://dx.doi.org/10.1007/978-3-032-26717-7_24.

Garrido Tenorio K, Becker-Peth M, de Vries J (2027) Human or AI Favoritism: Operational Impact of Automating Negotiations.

Guida M, Caniato F, Moretto A, Ronchi S (2023) The role of artificial intelligence in the procurement process: State of the art and research agenda. *Journal of Purchasing and Supply Management* 29(2):100823, ISSN 1478-4092, URL http://dx.doi.org/10.1016/j.pursup.2023.100823.

Herold S, Heller J, Rozemeijer F, Mahr D (2025) Brave New Procurement Deals: An Experimental Study of how Generative Artificial Intelligence Reshapes Buyer–Supplier Negotiations. *Journal of Purchasing and Supply Management* 101012, ISSN 14784092, URL http://dx.doi.org/10.1016/j.pursup.2025.101012.

Kesavan S, Kushwaha T (2020) Field Experiment on the Profit Implications of Merchants' Discretionary Power to Override Data-Driven Decision-Making Tools. *Management Science* 66(11).

Kirshner SN, Pan Y, Wu JX, Gould A (2026) Talking terms: Agent information in LLM supply chain bargaining. *Decision Sciences* 57(1):9–23, ISSN 1540-5915, URL http://dx.doi.org/10.1111/deci.70010.

Lei SI, Shen H, Ye S (2021) A comparison between chatbot and human service: customer perception and reuse intention. *International Journal of Contemporary Hospitality Management* 33(11):3977–3995, ISSN 0959-6119, 0959-6119, URL http://dx.doi.org/10.1108/IJCHM-12-2020-1399.

Logg JM, Minson JA, Moore DA (2019) Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes* 151:90–103, ISSN 07495978, URL http://dx.doi.org/10.1016/j.obhdp.2018.12.005.

Luo X, Tong S, Fang Z, Qu Z (2019) Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases. *Marketing Science* mksc.2019.1192, ISSN 0732-2399, 1526-548X, URL http://dx.doi.org/10.1287/mksc.2019.1192.

McManus S (2023) Can AI cut humans out of contract negotiations? URL https://www.bbc.com/news/business-67238386.

Revilla E, Saenz MJ (2026) How AI Is Reshaping Supplier Negotiations URL https://hbr.org/2025/07/how-ai-is-reshaping-supplier-negotiations.

Vaccaro MA, Caoson M, Ju H, Aral S, Curhan J (2025) Advancing AI Negotiations: New Theory and Evidence from a Large-Scale Autonomous Negotiations Competition. URL http://dx.doi.org/10.31219/osf.io/b3v9e_v1.

van Hoek R (2024) Insight from industry-early lessons learned about AI adoption in core procurement processes, directions for managers and researchers. *Supply Chain Management: An International Journal* 29(4):794–803, ISSN 1359-8546, 1359-8546, URL http://dx.doi.org/10.1108/SCM-02-2024-0143.

van Hoek R, DeWitt M, Johnson T (2022) How Walmart automated supplier negotiations. *Harvard Business Review* URL https://hbr.org/2022/11/how-walmart-automated-supplier-negotiations.

van Hoek R, Lacity M (2023) Procurement in the Age of Automation. *MIT Sloan Management Review* URL https://sloanreview.mit.edu/article/procurement-in-the-age-of-automation/.

von Schenk A, Klockmann V, Köbis N (2023) Social Preferences Toward Humans and Machines: A Systematic Experiment on the Role of Machine Payoffs. *Perspectives on Psychological Science* 17456916231194949, ISSN 1745-6916, 1745-6924, URL http://dx.doi.org/10.1177/17456916231194949.