



INTERNATIONAL JOURNAL OF HUMAN RIGHTS LAW REVIEW

An International Open Access Double Blind Peer Reviewed, Referred Journal

Volume 5 | Issue 5 | 2026

Art. 18

Who pays when the Machine pays? Legal Liability for Agentic UPI Transactions in India

Aprajita Sharma

LLM Student, 2nd Year,

Prof. Rajendra Singh (Rajju Bhaiya) University, Prayagraj

Recommended Citation

Aprajita Sharma, *Who pays when the Machine pays? Legal Liability for Agentic UPI Transactions in India*, 5 IJHRLR 197-208 (2026).

Available at www.ijhrlr.in/current-issues/.

This Article is brought to you for free and open access by the International Journal of Human Rights Law Review by an authorized Lex Assisto & Co. administrator.

For more information,
please contact submission.ijhrlr@gmail.com

Who pays when the Machine pays? Legal Liability for Agentic UPI Transactions in India

ABSTRACT

India's Unified Payments Interface (UPI) is entering a potentially significant transition from human-initiated digital payments to payments executed by software agents acting under pre-authorised instructions. In September 2026, reporting indicated that the National Payments Corporation of India (NPCI) was preparing a Unified Agent Protocol under which artificial-intelligence agents could execute limited UPI payments without requiring transaction-by-transaction approval. This development raises a legal question that existing payment regulation was not designed to answer: when a customer has authorised an AI agent but the agent subsequently makes an erroneous, excessive, manipulated or otherwise disputed payment, who bears the loss? This article examines whether India's existing framework – including the Payment and Settlement Systems Act, 2007, Reserve Bank of India directions on unauthorised electronic transactions, UPI mandate architecture and consumer-protection law – adequately allocates responsibility for agentic transactions. It argues that agentic payment disputes cannot simply be treated as either ordinary authorised transactions or conventional unauthorised fraud. They occupy an intermediate category in which human authority is delegated but machine judgement determines whether, when and to whom a payment is made. Using doctrinal legal analysis, regulatory analysis, recent empirical evidence on UPI scale and comparative consideration of emerging agentic-payment safeguards, the article proposes a layered liability model based on control, authentication, scope of delegation, auditability and system failure. It recommends revocable agent credentials, transaction and temporal limits, machine-readable mandates, mandatory audit trails, rapid suspension, consumer-first provisional redress and allocation of responsibility among banks, payment intermediaries, merchants and AI providers. The objective is to preserve the efficiency of agentic commerce while ensuring that automation does not dilute consumer protection.

KEYWORDS

Agentic payments, Artificial intelligence, Consumer protection, Liability, UPI

I. INTRODUCTION

Digital payments have transformed the legal and commercial meaning of authorisation. In a conventional UPI transaction, a customer initiates or confirms a payment through a payment application and the payment system processes that instruction. The model is comparatively straightforward: a human gives an instruction, authentication mechanisms validate the transaction, and the payment chain executes it.

Agentic commerce changes that sequence. An artificial-intelligence agent may be instructed to purchase goods whenever a condition is met, renew a service below a price ceiling, or make another routine payment without seeking a fresh approval for every transaction. The human still originates the authority, but software determines whether the conditions for payment have been satisfied. The central legal problem is whether existing concepts of authorisation and liability can accommodate a transaction in which human consent is transformed into bounded machine discretion.

Recent reporting indicates that NPCI is preparing a framework that could permit AI agents to make low-value UPI payments under customer-defined rules, spending limits, identity checks and audit trails. The reported proposal is expected to draw on existing mechanisms such as UPI Circle and Reserve Pay, while also contemplating a liability framework. These details remain under development and should not be treated as final law or regulation. Nevertheless, the proposal exposes a regulatory question before mass adoption occurs.

UPI's scale makes the question consequential. Reuters reported that UPI processed 24.51 billion transactions worth approximately 29.82 trillion in August 2026 alone. At such scale, even a small rate of agentic error or abuse could create a substantial aggregate consumer-protection problem. The legal architecture should therefore be designed before autonomous payment becomes normal rather than after disputes become widespread.

The article asks whether India's present framework adequately allocates liability when an AI agent acts within, outside or ambiguously beyond a customer's instructions. It argues that the answer is presently incomplete and that the emerging framework should recognise a distinct category of delegated machine action and allocate liability according to control over the relevant failure.

II. RESEARCH GAP AND LITERATURE REVIEW

Existing Indian legal scholarship on digital payments has primarily addressed electronic banking fraud, authentication, data protection, cybercrime, intermediary responsibility and consumer remedies. These bodies of law remain essential, but agentic payments introduce a different factual structure. A customer may intentionally grant an agent payment authority and yet dispute the particular transaction because the agent misunderstood the instruction, was manipulated by external information, exceeded a spending condition or acted on an erroneous interpretation.

The conceptual gap can be expressed as a three-part distinction: ordinary unauthorised payment, authorised delegated payment, and authorised delegation followed by disputed machine action. The third category is legally important because the customer has not necessarily lost control through classic credential theft, yet the customer may not have intended the particular payment. Treating every such transaction as fully authorised risks shifting too much risk to consumers; treating every machine error as unauthorised may undermine the purpose of delegation.

Recent policy commentary has begun identifying consent, authentication, cybersecurity, auditability and liability as principal governance questions for agentic UPI. However, the regulatory architecture remains emergent. This article therefore focuses on the liability problem at the boundary between payment regulation and machine agency.

III. RESEARCH METHODOLOGY

This study adopts a doctrinal and policy-oriented research design. Primary legal sources include the Payment and Settlement Systems Act, 2007, relevant Reserve Bank of India directions on customer liability for unauthorised electronic banking transactions, the Consumer Protection Act, 2019, and publicly available NPCI operational circulars concerning UPI mandates and delegated payment mechanisms. Secondary material includes recent reporting on India's proposed agentic-payment framework and international payment-network developments.

The scope is limited to consumer-facing UPI payments and principally concerns low-value or routine transactions for which a customer gives an AI agent advance authority. It does not attempt to resolve separate regulatory questions raised by autonomous investment decisions, lending, securities trading or high-value corporate payments.

The principal analytical framework is “liability follows control”. Under this approach, the participant best positioned to prevent a particular failure should ordinarily bear primary responsibility for it, subject to causation, contractual allocation and statutory protections. The framework is supplemented by consumer protection, proportionality, auditability and the principle that consumers should not have to reconstruct a complex technical chain before obtaining provisional redress.

IV. THE EMERGING AGENTIC UPI MODEL

Recent reports indicate that India is preparing a Unified Agent Protocol that could allow AI agents to execute small UPI payments without individual approval for every transaction. Reported safeguards include rule-based payment instructions, spending limits, identity checks and audit trails. The proposal is also reported to draw on UPI Circle, which permits delegated payment authority, and Reserve Pay, which can block funds for multiple debits. Because the protocol was still under development as of the date of this article, these features should be understood as reported proposals rather than settled regulatory requirements.

The legal significance lies in the transition from delegated payment authority to delegated decision-making. A human user can authorise another person to make a payment, but an AI agent may interpret a natural-language instruction and evaluate changing external conditions before deciding to transact. For example: “buy my regular groceries each week, but spend no more than 4,000”; “renew my broadband unless the price increases by more than 10%”; or “purchase the flight if the fare falls below 6,000”. Each instruction contains ambiguity that a machine must operationalise.

The agent therefore sits between intention and execution. A legal framework must be able to establish what instruction existed, what information the agent relied upon, what rule it applied, whether the transaction was within the authorised scope, and which participant controlled the relevant failure. Auditability becomes a substantive consumer-protection requirement rather than merely a technical feature.

V. INDIAN LEGAL FRAMEWORK AND ITS LIMITS

A. Payment and Settlement Systems Act, 2007. The Payment and Settlement Systems Act establishes the Reserve Bank of India as the authority for regulation and supervision of payment systems. It provides for authorisation of payment systems, regulatory standards, directions, supervision, and duties of system providers. This gives RBI

significant institutional capacity to prescribe technical and governance standards for emerging payment models. However, the Act is a framework statute for payment systems rather than a detailed liability code for autonomous software decisions.

B. RBI customer-liability framework. RBI's 2017 directions on unauthorised electronic banking transactions establish a customer-protection architecture under which liability depends on factors such as bank fault, third-party breach and customer negligence; importantly, the burden of proving customer liability lies on the bank. This is a strong starting point because payment security should not automatically place the entire loss on consumers. Yet agentic payments complicate the binary of authorised and unauthorised action. The customer may have authorised the agent's access while disputing the agent's particular decision. A future framework should expressly address this intermediate category.

C. UPI mandates and delegated authority. NPCI's UPI Autopay framework demonstrates that recurring digital payments can operate through mandates, while user controls include viewing, revoking, pausing and modifying mandates. Delegated-payment functionality through UPI Circle provides another institutional foundation. Nevertheless, an AI agent differs from a fixed recurring mandate because it may make contextual decisions about whether a transaction should occur.

D. Consumer Protection Act, 2019. The Consumer Protection Act provides a broad framework for protection of consumer interests and dispute resolution. Depending on the facts, deficiencies in service and unfair trade practices may become relevant. The challenge is procedural and substantive: consumers should not have to identify which layer of a multi-party agentic architecture failed before obtaining effective relief. A consumer-first redress mechanism can later allocate costs among responsible entities.

VI. THE CENTRAL LIABILITY PROBLEM: WHEN AUTHORISATION AND INTENTION DIVERGE

The most difficult cases are not obvious frauds. Consider four scenarios.

Scenario 1 – Agent exceeds a spending limit. A customer authorises an agent to spend up to 2,000 per transaction. Due to a system defect, the agent initiates a 7,500 payment and the payment system executes it. The customer's instruction is clear and the payment is outside the delegated scope. Liability should not ordinarily fall entirely on the customer merely because the customer initially authorised the agent.

Scenario 2 – Agent misinterprets an instruction. A customer asks the agent to purchase a refundable flight below 8,000. The agent identifies a 7,500 ticket but purchases a non-refundable fare. The payment may be technically within the price limit while being outside the substantive instruction. This illustrates why spending ceilings alone cannot define authorisation.

Scenario 3 – External manipulation. An agent visits a merchant website containing malicious instructions designed to redirect it to a different payment recipient. The customer authorised the shopping task but not the malicious redirection. Responsibility should turn on which participant failed to maintain agent identity, merchant authentication or transaction integrity.

Scenario 4 – Ambiguous instruction. A customer tells an agent to “buy the cheapest available option”. The agent purchases a product that is technically cheapest but materially different from the customer’s normal purchase. Here the dispute concerns interpretation rather than technical failure. The framework should require transparent records of the agent’s decision and, for higher-risk ambiguity, escalation to the human.

These scenarios show why the law should distinguish *delegation* from *blanket authorisation*. Delegation should create bounded authority. The machine should not become a substitute legal person with unlimited discretion; it should operate as a controlled digital delegate whose authority is defined by the principal and constrained by the payment system.

VII. LIABILITY SHOULD FOLLOW CONTROL

A workable framework should distribute responsibility according to the source of the failure.

Consumer: The consumer should bear losses caused by deliberate misuse, knowingly excessive delegation, failure to protect authentication credentials, or conduct that clearly violates the terms of the agent mandate. Consumer liability should nevertheless remain subject to statutory protections and reasonable limits.

Bank / payer-side payment service provider: Banks and payment providers should bear primary responsibility for failures in account controls, authentication, transaction-limit enforcement, credential management and mandated dispute handling. If a transaction clearly breaches a system-enforced ceiling but is processed nevertheless, the consumer should have a strong claim against the responsible payment participant.

AI-agent provider: Where the agent departs from an authenticated instruction because of defective software, inadequate safeguards, poor instruction handling, or failure to maintain the integrity of agent credentials, the AI provider should bear responsibility proportionate to its control over that risk. Providers should maintain sufficient logs to demonstrate how the agent arrived at a payment decision.

Merchant: Merchants should remain responsible for ordinary consumer-law duties, misrepresentation and fraudulent conduct. Where a merchant deliberately exploits an agent by presenting deceptive machine-targeted information or redirecting payment instructions, responsibility should not be shifted to the consumer merely because the consumer used an autonomous agent.

NPCI / system architecture: NPCI's role is different: as the operator of payment infrastructure and rule-setting environment, it should establish interoperability, credential, audit and dispute standards rather than become the default compensator for every downstream loss. System-level failures attributable to infrastructure design or rule implementation should, however, trigger appropriate regulatory and contractual responsibility.

VIII. COMPARATIVE AND INTERNATIONAL PERSPECTIVE

Agentic payments are not an exclusively Indian development. International card networks and payment providers are also developing infrastructure intended to let AI agents transact while preserving controls over authentication, identity and visibility. Recent reporting indicates that Mastercard and Visa are developing agentic capabilities for commerce and that Mastercard has demonstrated authenticated agentic transactions in India. These developments suggest that liability and authentication are emerging as global design questions.

India nevertheless has a distinctive advantage: UPI already combines interoperability, mandate infrastructure and a centralised rule-setting ecosystem. Rather than importing a foreign model wholesale, India can use these characteristics to establish a domestic standard for machine-mediated payment authority.

The most useful comparative principle is not the creation of a separate legal personality for AI. Agentic payments should instead be treated as a form of bounded delegation. The agent's technical autonomy should not erase the legal chain of responsibility connecting the consumer's mandate, the agent's credentials, the payment provider's controls and the merchant's conduct.

IX. RECOMMENDATIONS

1. **Create a regulatory category of “agent-initiated payment”:** RBI and NPCI should expressly distinguish agent-initiated transactions from ordinary user-approved payments and conventional unauthorised transactions.
2. **Make mandates machine-readable and scope-specific:** A valid agent mandate should specify transaction ceiling, aggregate daily/monthly ceiling, merchant or category restrictions where relevant, geographical scope where relevant, expiry, revocation mechanism and escalation triggers.
3. **Require attributable agent credentials:** Every agent should have an identifiable credential linked to the customer mandate and the responsible agent provider. Payment records should allow investigators to determine which agent initiated a transaction.
4. **Require mandatory audit trails:** Each agentic payment should record the relevant user instruction, applicable rule, agent identity, merchant identity, amount, timestamp and reason the transaction was considered permissible. Sensitive internal model data need not automatically be exposed to consumers, but enough evidence should exist for dispute resolution.
5. **Introduce consumer-first provisional redress:** Where a customer promptly disputes an agentic transaction, the relevant regulated entity should provide rapid provisional relief, especially where the transaction clearly exceeds an authenticated limit or violates an enforceable system rule.
6. **Allocate liability according to control:** Regulations should establish rebuttable presumptions: banks/payment providers for authentication and system-control failures; AI providers for agent-behaviour failures; merchants for deceptive or fraudulent conduct; and consumers only where their own conduct materially caused the loss.
7. **Impose a kill-switch and revocation mechanism:** Consumers should be able to immediately suspend an agent through a simple channel. Revocation should propagate across participating payment providers rather than requiring consumers to contact multiple entities.
8. **Use risk-tiering:** Low-value routine purchases may operate under pre-authorisation. Higher-risk transactions – financial investments, unusually large purchases, new beneficiaries,

unusual merchants or transactions outside established parameters – should trigger human confirmation.

9. **Use regulatory sandboxes before broad deployment:** Agentic UPI should initially be tested in controlled environments with incident reporting, security testing, consumer complaints data and independent audits.
10. **Keep consumers outside technical blame allocation:** The consumer should have a single accessible dispute channel. Banks, PSPs, AI providers and merchants can subsequently determine contribution between themselves.

X.FINDINGS

The study yields five principal findings. First, agentic payments create a legally distinct problem because human delegation and machine decision-making can diverge. Second, India's existing payment framework contains useful building blocks – mandates, delegated authority, system regulation and customer-liability rules – but does not yet provide a complete liability architecture for autonomous machine action. Third, UPI's enormous transaction scale makes early governance particularly important. Fourth, liability should follow control over the relevant risk rather than defaulting to either the consumer or the payment system as a whole. Fifth, auditability and rapid revocation are as important as authentication because a payment can be authenticated as originating from an authorised agent while still being inconsistent with the consumer's substantive instruction.

XI.CONCLUSION

Agentic payments represent a change in the legal structure of digital commerce, not merely another payment-interface innovation. When a customer authorises an AI agent to act under rules, the transaction chain becomes: human intention → machine interpretation → agent decision → payment authentication → settlement → merchant performance. A dispute may arise at any point in that chain.

India has an opportunity to address the problem before agentic commerce becomes ubiquitous. Its existing UPI architecture, combined with RBI's consumer-protection framework and NPCI's mandate infrastructure, provides a strong foundation. What is required is an additional governance layer that converts human intent into bounded, auditable and revocable machine authority.

The core legal principle should be simple: the AI may spend, but the

human remains the source of authority – and the entity that controls the relevant failure should ordinarily bear the resulting responsibility. If India adopts this principle through clear rules on agent identity, scope, limits, auditability, redress and risk-tiering, agentic UPI can develop without sacrificing the trust on which digital payments depend.

ENDNOTES AND REFERENCE

1. Reuters, “India preparing rollout of agentic payments on UPI, sources say”, 1 September 2026. The report states that India is developing a framework for AI agents to conduct small UPI payments without transaction-by-transaction approval and reports proposed safeguards including spending limits, identity checks, audit trails and liability provisions.
2. Reuters, *ibid.* The report states that UPI processed 24.51 billion transactions worth approximately 29.82 trillion in August 2026.
3. The Payment and Settlement Systems Act, 2007, Act No. 51 of 2007, India Code. The Act provides for regulation and supervision of payment systems and designates RBI as the authority for that purpose.
4. Reserve Bank of India, “Customer Protection – Limiting Liability of Customers in Unauthorised Electronic Banking Transactions”, RBI/2017-18/15, DBR.No.Leg.BC.78/09.07.005/2017-18, 6 July 2017. The framework addresses customer liability and places the burden of proving customer liability on the bank.
5. National Payments Corporation of India, UPI Operational Circular No. 223/2025-26, “Enhancement of UPI Autopay”, 7 October 2025. The circular provides for lifecycle management of UPI Autopay mandates, including viewing, revocation, pausing, modification and user-driven portability.
6. National Payments Corporation of India, publicly available UPI documentation concerning delegated payments/UPI Circle. The precise circular/version relied upon should be checked against the latest NPCI publication available on the date of submission.
7. Consumer Protection Act, 2019, Act No. 35 of 2019, India Code. The Act provides for protection of consumer interests and mechanisms for administration and settlement of consumer disputes.
8. ETGovernment, “Who pays when AI gets it wrong? The governance challenge of agentic payments on UPI”, 3 September

2026. The article discusses delegated payment, delegated judgement, consent, auditability and liability.

9. ETBFSI, “Will proposed agentic UPI raise new fraud, authentication risks for banks?”, 2 September 2026. The report discusses authentication, fraud-monitoring and dispute-resolution implications.