The Dialogue®
INFORM ENGAGE IDEATE

SUBMISSION OF COMMENTS

on

Reserve Bank of India's Discussion Paper on **Exploring Safeguards** in Digital Payments to Curb Frauds

DATE OF DRAFT:
April 9, 2026

SUBMITTED ON:
May 8, 2026

This submission has been informed by consultations with subject matter experts, industry leaders, former law enforcement personnel, and researchers.

SUBMISSION OF COMMENTS ON RESERVE BANK OF INDIA'S DISCUSSION PAPER ON EXPLORING SAFEGUARDS IN DIGITAL PAYMENTS TO CURB FRAUDS

Authors: FinTech and Sustainable Finance Team.

The Dialogue is a public policy think tank with a vision to drive a progressive narrative in India's policy discourse. Founded in 2017, we believe in facilitating well-researched policy debates at various levels to help develop a more informed citizenry, on areas around technology and development issues. The Dialogue has been ranked as the world's Top 10 think tanks to watch out for, by the Think Tank and Civil Societies Programme (TTCSP), University of Pennsylvania in their 2020 and 2021 rankings.

For more information
visit thedialogue.org.in

Publication date: July 2, 2026

Catalogue no.: TD/DE/WC/0726/10

Suggested Citation: FinTech and Sustainable Finance Team. (2026). Submission of Comments on the Reserve Bank of India's Discussion Paper on Exploring Safeguards in Digital Payments to Curb Frauds. The Dialogue.

Contents

A. Introduction

The Reserve Bank of India (RBI) released a discussion paper titled *Exploring Safeguards in Digital Payments to Curb Frauds* in April 2026, which proposes four interventions to address Authorised Push Payment (APP) fraud and related risks in the digital payments ecosystem.

With regard to this, The Dialogue is pleased to submit its comments on the discussion paper. Drawing upon our extensive research in this area and active engagement with the policy ecosystem, this submission incorporates insights from a diverse group of stakeholders, including subject matter experts, industry leaders, former law enforcement personnel, and researchers. It consolidates the technical, legal, and operational views that emerged from our consultations to provide a comprehensive perspective on the proposed framework.

B. Observations

- The fraud landscape has shifted from technical compromise to social engineering. Basically, account takeovers are now a small share of reported cases, and most losses come from APP fraud where the victim is manipulated into authorising the transaction.
- The Indian payments architecture was built for speed, and a lot of the financial inclusion gains of the past decade rest on this design choice. This was built on the strong Telcom infrastructure,
- which is closely integrated with the financial sector now. Any law enforcement action to introduce friction in this architecture to curb fraud risks reverses those gains, particularly for first-time users and lower-income segments. The general consensus is tilted in favour of legitimate friction, applied to genuinely suspicious transactions, over blanket friction. This can be done using intelligence generated by technology platforms outside the banking perimeter, such as telecom network signals.
- The measures proposed in the discussion paper should be piloted in a controlled sandbox before a mandatory national rollout. A major reason for this is that the interaction effects between four simultaneous interventions are difficult to model on paper. A phased approach would surface implementation issues before they affect the entire ecosystem.

1: Lagged Credit for Authorised Push Payments

The proposal applies to a mandatory one-hour lag at the payer's bank for APP transactions above ₹10,000, with the option to whitelist payees or specific transactions. While the intent to break the psychological pressure by the fraudsters is logical, the design has a few concerns.

Comments

- The ₹10,000 threshold is a blunt limit, i.e., it captures a large share of fraud cases by volume but also captures most legitimate person-to-person transfers that Indian users make routinely.
- The most likely behavioural response from fraudsters to this friction would be transaction splitting, basically, breaking a larger sum into multiple sub Rs. 10,000 transfers that bypass the same. Moreover, rigid friction for all high-value transfers might shift users toward less secure informal channels or cash, creating second-order economic effects. This pattern has been observed in other jurisdictions with similar thresholds, and there is little reason to expect a different pattern in India.
- It is essential to clarify the intention that this proposal seeks to resolve and whether it targets P2P or P2M flows. Notably, B2B transactions are generally not part of these retail-centric fraud patterns and should remain excluded to avoid disrupting commercial liquidity.
- Unlike the NPCI ecosystem, legacy systems such as IMPS and RTGS often lack sophisticated, real-time fraud prevention and signalling systems. Imposing blanket lag on one system while others remain vulnerable may simply shift the fraud vector to less-monitored channels.
- In specific fraud cases, such as Digital Arrest scams, victims are often under such intense psychological coercion that they deliberately navigate friction journeys, such as adding new payees or waiting out cooling periods, to complete the transaction. Unlike impulsive UPI transfers, these victims are often "convinced" to follow through, meaning a time lag alone may not be a sufficient deterrent.

- Given current transaction patterns, this proposal may not impact a large number of people in terms of total transaction volume, but it imposes a uniform cost on all users irrespective of their actual risk profile.

Recommendations:

- A fixed one-hour delay also imposes a uniform cost on all users, irrespective of the actual risk profile of the transaction. A more targeted approach would use UX based interventions such as high-severity warning screens, mandatory cooling prompts for new payees, or a confirmation step for transactions that deviate from the user's historical pattern.
- Instead of a lag at the payer's end, propose that funds be blocked at the payee's end (Shadow Credit). This keeps the payer's experience "instant" while allowing a 60-minute window for reversal if fraud is reported.
- At the point of account onboarding and underwriting, non-banking behavioural signals, including mobile identity history, device patterns, and location consistency, offer a live verification dimension that is significantly harder for fraudsters to fabricate than document-based KYC alone. This is particularly relevant for detecting mule accounts, where paper credentials may appear legitimate but network behaviour signals elevated risk.
- An architecture that exempts or whitelists contacts stored on the phone from this time lag can also be explored. For instance, recommend exempting any P2P transaction where the payer has a successful, non-disputed history with the payee within the last 12 months.
- Suggest that while ₹10,000 is the default, users should be empowered to customise their own threshold (e.g., ₹20k or ₹50k) in real-time, similar to card settings.
- Lastly, user literacy is a focal point as well. The end user needs to understand the relevance of this time lag and not perceive it as a mild annoyance. Without deep literacy, users may simply wait out the hour and proceed with the transfer under the fraudster's continued guidance.

2: Additional Authentication by a Trusted Person for Vulnerable Sections

The proposal mandates an additional authenticator for citizens aged 70 and above and Persons with Disabilities (PwD), for transactions above ₹50,000.

The underlying concern that certain user segments are disproportionately affected by APP fraud is valid, but, the proposed design of this intervention presents practical and regulatory challenges.

Comments

- Vulnerability to APP fraud is situational, and prevailing research highlights that demographics might not be a considerable factor. Additionally, recent case patterns show that even younger users have been compromised during digital arrest scenarios involving sustained psychological pressure. Defining vulnerability purely by age or disability status underprotects users outside these categories and overprotects users within them, a lot of whom transact independently and competently. The definition of 'vulnerable' needs to be more inclusive, while remaining rooted in dynamic risk-based profiling rather than static demographic markers.
- The trusted person model has its own set of problems because the authenticator has neither legal nor beneficial interest in the account, but effectively gates outward transactions. This raises questions about liability when the authenticator approves a fraudulent transaction or refuses a legitimate one. It also introduces delays where the trusted person is not immediately available, particularly for older users whose trusted contacts may themselves be elderly.
- A complicated "trusted person" framework might inadvertently encourage new types of fraud, where scammers socially engineer themselves into becoming the designated "trusted person" for a senior citizen.
- Worth noting that fraud is migrating toward mandates, where the initial small debit bypasses protection, but subsequent high-value debits occur without user awareness. The ₹50,000 threshold should apply to the mandate's maximum limit rather than just the first debit.
- Additionally, a proper legal framework should define the responsibilities and liability of the trusted person in a transaction, since the role currently sits outside any defined banking relationship.

Recommendations

- Efforts should be directed at the implementation of passive fraud detection measures at the application and network levels. For instance, at the telecom level, there are sufficient signals to flag fraud before the consumer is even reached. These infrastructure-level flags provide a stronger defence than manual third-party intervention.
- A more effective alternative is behavioural AI as the digital guardian. The bank's fraud detection system can monitor transactions against the user's historical pattern, and when a transaction deviates significantly, the system can intervene with additional authentication or a cooling prompt. This approach does not depend on the availability of a third party and adapts to the actual fraud patterns observed in the user's account. This can include silent authentication through SIM-swap detection APIs and other telecom network signals such as mobile identity events, device behaviour, communication pattern anomalies, and location data; as well as device fingerprinting, behavioural biometrics, and cross-platform identity signals. Since these signals are generated continuously, they can be observed hours before a fraudulent transaction is attempted.
- Recommend shifting from a mandatory age-based system to an opt-in safety system. This respects the independence of capable seniors and PwDs while protecting those who self-identify as needing support
- Shift from static demographic markers to a more inclusive, risk-based approach that covers various vulnerable groups, while functioning as a last line of defence.

3: Accounts to Receive Credits Commensurate with the Nature of Relationship

The proposal caps annual aggregate credits at Rs. 25 lakh for low credit turnover accounts, with shadow credits beyond this limit pending additional documentation. This is the most operationally heavy of the four proposals and runs into a fundamental conflict with parallel financial inclusion efforts.

Comments

- A lot of the recent push in Indian financial services has been towards using alternate data, i.e., account flows, income receipts, cash management patterns, etc., to extend credit to MSMEs and gig workers who lack traditional documentation. The proposed threshold cuts across this since a gig worker with multiple income sources, or a small traders/MSME whose receipts spike during festival season, may breach the Rs. 25 lakh ceiling without any underlying fraud. In such cases, the shadow credit and 30-day documentation window becomes a friction point that disproportionately affects the segments the rest of the policy ecosystem is trying to serve.
- Also, if breaching the threshold requires producing documents at the home branch to upgrade the limit, it tends to be exclusionary and a huge cost burden for the migrant workforce.
- The compliance burden on banks is also high as banks will have to track aggregate credits across the year for every individual, decide flag status at onboarding and on an ongoing basis, and manage the shadow credit process.
- Adding to the above, the discussion paper does not specify what variables determine the Nature of Relationship that justifies turning the flag off, which creates implementation uncertainty for banks and inconsistent outcomes for customers.
- Measures introduced today must be technologically and operationally sensitive so they do not inadvertently create new regulatory or accessibility issues in the future. Over-reliance on static manual documentation is a regressive step in an increasingly digital economy.

Recommendations

- Before a national-level rollout, RBI can lead a collaborative pilot exercise involving all industry stakeholders (Banks, Fintechs, and TSPs). This pilot could focus on identifying and reducing "outliers", legitimate users who may be unfairly flagged by the ₹25 lakh limit, to ensure the final threshold is evidence-based and minimises collateral damage to the digital economy.
- Recommend that non-revocable mandates (such as those for IPOs or loans) should be exempted from the ₹25 lakh cap, as these involve legal commitments to pay.
- A more proportionate approach would use real-time risk scoring at the account level, drawing on data from the Financial Intelligence Unit (FIU), Digital Intelligence Unit (DIU), telecom networks, and the

Digital Payment Intelligence Platform (DPIP) that the RBI is already building with the Reserve Bank Innovation Hub (RBIH). Hence, the RBI should consider whether the additional layer proposed here adds enough fraud prevention value to justify the operational and inclusion costs.

- Focus could be on non-intrusive, technical measures that flag suspicious velocity or source of fund patterns without requiring the users to intervene, unless a high-confidence fraud signal is triggered. The payee side signals can be supported with telecom intelligence to detect mule account behaviour of the receiving party as part of the credit assessment workflow.
- If the proposal proceeds, the threshold should be calibrated by account type and behavioural history.
- The payee side signals can be supported with telecom intelligence to detect mule account behaviour of the receiving party as part of the credit assessment workflow.

4: Customer Induced Controls

The proposal extends customer controls and a kill switch facility across all digital payment channels, with the option of disabling digital payments by default for new accounts. However, there are a couple of issues with the same:

Comments

- Efforts should be prioritised toward digital routing, ensuring that the trail of funds through inactive or "mule" accounts is highly traceable. Rather than simply blocking reactivation, the focus should be on how these accounts are routed within the larger ecosystem to identify laundering hubs.
- The kill switch is, by its nature, a reactive control. The main reason for this is that by the time a customer realises the account has been compromised and triggers the kill switch, funds have often already been moved through the mule account network or withdrawn at ATMs. Its primary value lies in containing further loss rather than preventing the initial fraud.
- Additionally, within this, the reactivation route is problematic. If the only way to reactivate digital payments after a kill switch is a physical branch visit, the measure risks excluding migrant workers, rural users, and customers whose home branch is far from their current location. For a lot of customers, the digital channel is the only practical route to their account. A digital reactivation pathway, with stringent authentication, should be the standard. The branch visit can remain available as one option among many, not the only one.
- Rigid thresholds on reactivated accounts can be easily bypassed by fraudsters who "warm up" accounts with small transactions before executing a scam.

Recommendations

- The payee side signals can be supported with telecom intelligence to detect mule account behaviour of the receiving party as part of the credit assessment workflow.
- For new accounts where limited behavioural analytics are available, a mandatory one-hour lag for high-value transactions could be implemented as a temporary safeguard until a baseline behaviour profile is established.
- Prepaid Payment Instruments (PPIs) should be segregated into limited value categories. By capping the stored value and transaction limits for non-KYC or basic PPIs, the systemic risk of these instruments being used for high-value fraud is naturally mitigated.
- Suggest that the "kill-switch" should specifically apply to revocable mandates. This addresses the issue where users currently feel they have no control over fraudulent merchants and resort to "emptying their accounts" as a makeshift kill switch.
- Beyond mandates and standing instructions, Government benefit transfers (DBT) and emergency medical payments should be exempted from the kill-switch to prevent life-critical exclusions.
- Default off for new accounts is also operationally challenging since a lot of new account openings in India happen alongside the customer's first exposure to digital payments. Disabling digital channels by default creates an additional onboarding step that may be poorly understood and may push customers to the cash channel instead. A graduated activation approach is more practical, where low-risk channels such as inward credits and merchant payments are enabled by default, while higher-risk functions such as outward P2P transfers above a small threshold or International Roaming/Transactions require explicit

opt-in. Further, suggest that digital reactivation could only be possible from a "Trusted Device" already known to and verified by the bank.

- The controls and the kill switch also need to be available in regional languages with simple visual cues, contrary to dense legal disclaimers. A control that the customer cannot easily understand or operate fails its purpose. Similarly, instead of automated blocks, reactivation should trigger mandatory, high-impact awareness and education prompts within the app. This ensures the user is informed of current fraud trends before resuming high-frequency transacting.

C. Concluding Remarks

Most of the proposals currently recommended focus on downstream interventions, measures that occur at the final point of transaction. However, our extensive research and stakeholder interactions on digital fraud typology reveal that these are often the easiest for fraudsters to detect, predict, and break through evolving tactics like transaction splitting or psychological grooming.

Moving forward, the policy focus must shift from restrictive consumer controls to fraudster behaviour detection. A robust defence-in-depth strategy should prioritise behavioural analytics, transaction velocity, and pattern recognition at the infrastructure level.

1. Policy Level Recommendations

- **Shift to Behavioural Analytics:** Regulations need to move away from static, blanket thresholds and toward mandates for real-time behavioural analytics and pattern recognition.
- **Protection of Vulnerable Customers:** India can look to adopt proactive institutional frameworks similar to the **UK's Financial Conduct Authority (FCA)** directives on the fair treatment of vulnerable customers. This ensures that "vulnerability" is treated as a situational risk rather than a demographic burden.
- Incentivising **Passive Defence:** Policy should prioritise "silent" infrastructure-level defences over active consumer interventions to maintain the ease of use that characterises India's digital payment success.

2. Institutional Level Recommendations

- **Cross-Industry Intelligence Layer:** A formal mechanism must be established to share real-time intelligence between Telecom Service Providers (TSPs), banks, and digital platforms to intercept fraud signals before they reach the user. The Digital Payment Intelligence Platform (DPIP), currently under development by the Reserve Bank Innovation Hub (RBIH), can serve as the central node for risk-scoring and account-level monitoring. Reserve Bank of India's Digital Payment Intelligence Platform (DPIP), which is intended to generate real-time, transaction-level risk scores based on beneficiary profiles, currently relies predominantly on banking transaction data. There is a strong opportunity to evolve DPIP into a truly multi-source intelligence engine—one that incorporates non-banking signals, including telecom network data, device intelligence, and behavioural analytics, alongside banking data and Indian Cyber Crime Coordination Centre (I4C) fraud reports.
- **Industry-Wide Pilot Exercises:** Before finalising national mandates, the RBI should lead collaborative pilots with all stakeholders to identify "outliers" and calibrate thresholds based on empirical evidence rather than assumptions. We recommend that the RBI formally pilot industry solutions bringing together banks, TSPs, NPCI, and technology providers to test and validate fraud prevention solutions in a controlled setting before full-scale deployment. This approach would also allow for fine-tuning of risk thresholds that trigger friction, helping to balance security requirements with customer convenience.
- **Recall Request Protocol:** Propose a standardised, interoperable protocol where a remitting bank can send a "Fraud Alert" signal to a beneficiary bank. The receiving bank should then be mandated to instantly freeze the equivalent amount for 24 hours, even if the funds have already been moved to a secondary account.
- User-Defined Custom Limits: Recommend that all users (not just vulnerable ones) be given the ability to set custom limits across different categories, like P2P and P2M transactions.

thedialogue.org.in

LinkedIn | The Dialogue

X | The Dialogue

Whatsapp | The Dialogue

Instagram | The Dialogue