Who Pays for Payment Fraud? Optimal Detection and Liability Rules

Vimal Balasubramaniam Thomas Mosk Antoine Uettwiller*

May 1, 2025

Abstract

We develop a dynamic model of payment fraud detection in which fraudsters adapt to detection technologies by shifting to harder-to-detect methods, leading to a decline in model performance over time. We derive the socially optimal level of fraud screening and show that full liability induces efficient investment by a monopolist payment service provider (PSP). However, this liability rule is not always optimal, especiall in competitive settings. In competitive markets, full liability leads to excessive screening, as each PSP's effort intensifies fraudster adaptation and imposes a negative externality on others. Full liability also reduces consumer vigilance, which supports the use of partial reimbursement and a consumer excess policy design. Finally, we show that the optimal allocation of liability between sending and receiving PSPs depends on the relative effectiveness of the receiver's know-your-customer procedures and the sender's transaction monitoring systems. Our findings highlight the importance of incorporating dynamic fraudster adaptation in designs of payment system policies.

JEL Classification: G21, G23, G28, L51

*For useful suggestions we thank Jason Sturgess, Radoslawa Nikolowa, and seminar participants at the Queen Mary University of London. Balasubramaniam: Queen Mary University of London, CEPR. Email: v.balasubramaniam@qmul.ac.uk. Mosk: Queen Mary University of London. Email: t.mosk@qmul.ac.uk. Uettwiller: Queen Mary University of London. Email: a.uettwiller@qmul.ac.uk.

1 Introduction

Payment fraud exploits consumers through impersonation, fake investment, and invoice scams, extracting billions annually. In 2024 alone, U.S. consumers lost $12.5 billion to payment fraud, U.K. consumers 341 million, EEA consumers 4.3 billion, Indian consumers 14.5 billion ($ 175 million). Modern digital payment systems such as Brazil's instant payment system PIX, or India's Unified Payment Interface (UPI) have on boarded hundreds of millions of new customers who are potentially more vulnerable to fraudsters, making this a salient and important global challenge. Payment scams are often difficult and costly to detect for individual consumers, whereas payment service providers (PSPs) can exploit scale economies and machine-learning techniques to detect fraud more efficiently. Yet, when consumers bear most fraud losses, PSPs have little incentive to invest in detection. To correct this, regulators have shifted liability onto PSPs—e.g. the U.K. PSR's reimbursement rules and the EU's 2023 Payment Services Proposal—thereby stimulating enhanced fraud-detection. Other countries such as the U.S., and India, are under the consumer-bears-it-all regime. The question of how these policies affect welfare, and the optimal design of the balance between incentives to invest in detection technologies for PSPs, and consumer vigilance therefore becomes relevant.

This paper examines how liability rules shape fraud prevention in digital payments. We focus on "authorized push payment" (APP) fraud, where consumers are tricked into sending money to scammers. First, we ask how liability allocation affect PSPs' investment in fraud detection. We then evaluate how competition and consumer moral hazard (the flip side of consumer vigilance) affect optimal liability policies. Finally, we ask how policymakers can design rules that mitigate the negative side-effects of shifting payment fraud liabilities to PSPs.

Unlike unauthorized transactions, such as stolen credit card information, APP fraud occurs when criminals deceive victims into willingly authorizing payments to fraudulent accounts. For example, TIME Magazine reported on a customer who was defrauded of $1,000 by a caller impersonating a Wells Fargo representative. Although the customer notified Wells Fargo immediately after realizing the funds were stolen, the bank initially refused to reimburse, arguing that under federal rules it was only required to cover truly "unauthorized" transactions—those made without any customer initiation—and that this impersonation scam, which induced the customer to authorize the transfer, did not qualify. Only after media scrutiny did Wells Fargo voluntarily refund the $1,000.[1] Such stories are

[1]TIME, 2024, Banks Aren't Doing Enough to Protect Customers From Scams, source: `https://time.com/6952817/financial-scams-banks`

not uncommon in other parts of the world. Since becoming two of the largest payment systems in the world, the scams on the Indian UPI and Brazil's PIX have risen.[2]

At first glance, shifting liability to PSPs seems an obvious solution – making PSPs bear the cost should incentivize efficient detection. This response parallels a central result in banking theory: banks provide unique services in the form of publicly unobservable screening and monitoring of borrowers, and for them to have the incentive to provide an efficient level of these services, it is necessary that they retain part of the loans they originate [Pennacchi, 1988, Gorton and Pennacchi, 1995, Parlour and Plantin, 2008, Gryglewicz et al., 2024]. However, this ignores two critical dynamics that are important while considering payment systems.

First, sophisticated fraudsters adapt their methods when faced with screening. As PSPs intensify detection of prevalent scams (e.g. fake invoices), fraudsters substitute toward harder-to-detect techniques (e.g. AI-generated impersonations).[3] We label this the "whack-a-mole" effect, where increased screening unintentionally accelerates fraud evolution. Because screening effort shifts the distribution of fraud types, the underlying data-generating process changes—a phenomenon known as "concept drift" in the computer-science literature, where the model performance decays over time [Talukder et al., 2024, Hernandez Aros et al., 2024, Chatterjee et al., 2024]. This finding departs from standard credit screening models that assume fixed detection accuracy (e.g. Boyd and Prescott, 1986, Gorton and Pennacchi, 1995). We therefore extend these models by directly modeling detection accuracy – accuracy declines in response to higher PSP screening effort as fraudsters adapt.

Second, competition in the payment system intensifies model decay by creating a negative screening externality: each PSP optimizes its own detection effort without accounting for the market-wide impact.[4] Fraudsters adapt to the average screening intensity – so even as an individual PSP captures its private gain from thwarted attacks, it ignores the collective cost of pushing scammers toward undetectable methods. Consequently, equilibrium screening effort rises with the number of PSPs. This is different from the [Hauswald and Marquez, 2006] result where competition decreases screening investment because of lower intermediaries' rents.

Payment-fraud screening is under provided when PSPs do not bear fraud losses, yet under full liability PSPs overinvest in fraud screening in competitive markets. In a monopoly,

[2]See, for instance, `https://www.bbc.co.uk/news/articles/c288m1km01po`, and for how policymakers are responding to it, here: `https://paymentscmi.com/insights/pix-scams-role-of-banks/`.

[3]For example, in India we see: `https://indianexpress.com/article/technology/he-downloaded-a-whatsapp-image-minutes-later-rs-2-lakh-was-gone-9951435/`.

[4]Brazil's PIX is a state-owned enterprise, thus an exception to this observation about competition.

we demonstrate that full liability – making the PSP internalize all fraud costs – correctly aligns private and social incentives, delivering first-best screening and full consumer reimbursement. In competitive payment systems, however, full liability leads to excessive screening due to negative spillovers. The socially optimal rule instead assigns partial liability, balancing incentives to screen against the intertemporal distortion created by dynamic fraud behavior.

One immediate concern is that full liability policy makes PSPs accountable for consumers' negligence, leading to moral hazard and diminished consumer vigilance – such as recognizing suspicious communications, verifying request authenticity, and exercising caution in financial transactions. Introducing consumer moral hazard forces the regulator to make a trade-off between incentivizing PSPs to screen and preventing reduced consumer vigilance. Customers might become less vigilant about protecting themselves from scams, knowing they will automatically get reimbursed. We show that the optimal liability rule is a partial liability rule in which both consumers and PSPs bear the cost of the APP scam fraud. In a more competitive payment market, the regulator optimally sets a lower PSP liability to prevents PSPs from over-investing in fraud screening technologies.

We then turn to understanding two important real-world implication of our dynamic model of fraud prevention. What happens when PSPs close accounts they suspect of fraud involvement? Who bears the cost of a realized payment fraud?

PSPs have the discretion to close accounts when they suspect fraud involvement, a policy intended to deter malicious behavior and protect the integrity of the system. However, when applied too aggressively or based on flawed signals – such as inaccurate fraud markers – this mechanism can lead to excessive account closures, harming innocent users and potentially deterring legitimate financial activity. In 2023, the consumer group *Which?* in the U.K. raised alarms over U.K. banks erroneously flagging customers as fraud risks, leading to unwarranted account terminations. The scale of the issue is underscored by the Financial Ombudsman Service, which reported over 1,380 complaints about current account closures in 2022–23. This institutional concern maps closely to the mechanism in our model, where account closures—typically triggered by successful fraud—disproportionately remove high-risk users from the platform, altering the population composition and feeding back into future screening and fraud dynamics. We show that the welfare implications of account closures are ambiguous. On the one hand, they weaken *ex ante* screening incentives by acting as a substitute: if risky users can be purged after fraud is realized, the marginal benefit of investing in early detection falls. On the other hand, when consumer fees are higher, debanking becomes more costly for PSPs, which strengthens incentives to screen more aggressively up front. This generates a wel-

fare trade-off: higher fees raise screening effort and reduce the number of users removed from the system, but they also increase the cost of access to the payment service.

Fraud prevention in payments hinges on two actors: the sending PSP, which screens transactions (e.g., using machine learning to flag suspicious transfers to new payees), and the receiving PSP, which vets account holders (e.g., use know-your-customer (KYC) checks to detect money mule accounts). Their efforts are interdependent – when one invests more in screening, the other tends to invest less, creating a free-rider problem. This leads to underinvestment overall, as neither fully internalizes the system-wide benefits of fraud detection. Optimal liability rules must account for this tension. For example, placing too much responsibility on sending PSPs can weaken incentives for receiving PSPs to scrutinize fraudulent accounts. Our work suggests that regulators should assign greater liability to receiving PSPs, as their role in preventing money mule accounts is often more critical, especially in an environment with high decay in returns to detection technology.

The allocation of liability critically depends on the effectiveness of the sending PSP's fraud detection. When transaction screening models degrade rapidly—as fraudsters adapt to evade detection—assigning too much liability to the sending PSP becomes counterproductive. In such cases, heightened screening efforts by senders accelerate the obsolescence of detection tools, while receiving PSPs underinvest in vetting fraudulent accounts. This dynamic pushes regulators toward a liability rule that leans more heavily on receiving PSPs, whose know-your-customer (KYC) checks face less immediate decay. The U.K's 50/50 liability split, while a step forward, may still fall short of the welfare-maximizing balance.

Our paper contributes to the literature on fraud in financial markets. While much of the existing literature examines the perpetrators of fraud—such as financial advisors [Dimmock et al., 2018, Dimmock and Gerken, 2012, Egan et al., 2019], CEOs [Khanna et al., 2015, Agrawal et al., 1999], and firms [Piskorski et al., 2015, Povel et al., 2007, Dyck et al., 2010, 2024]. This paper shifts the focus to the role of payment service providers (PSPs) in fraud prevention and compensation policies, by analyzing how PSPs' incentives shape fraud screening intensity. We show that when PSPs do not bear fraud costs, they tend to underinvest in fraud detection relative to the social optimum. However, mandatory reimbursement policies can correct this misalignment by forcing PSPs to internalize fraud risks, thereby improving screening incentives. At the same time, optimal compensation policies must consider consumer moral hazard.

Our research is related to the literature on information technology investment in financial markets [Hauswald and Marquez, 2006, Vives and Ye, 2025]. Hauswald and Marquez [2006] show that banks invest in screening technology to 'steal business' from other banks,

which could result in over-investment in screening in equilibrium. However, with more banks, the investment in screening technology for each bank falls because more competition reduces the returns to information acquisition. While the existing literature focuses on credit markets, our paper is the first paper studying investment in screening technologies in the payment markets. Parlour et al. [2022] studies FinTech competition in payment markets, but links this leads to the erosion of informational rents of banks in the credit market and the value of data portability.

2 The Model

We analyze a market for payment services in which transactions are vulnerable to targeted fraud attempts. In these scams, fraudsters deceive consumers through social engineering tactics—posing as bank representatives, romantic partners, investors, or legitimate businesses—in order to manipulate them into authorizing a payment. The funds are typically routed to accounts controlled by the fraudster, often via money mules who help obscure the money trail. This type of fraud is known as Authorized Push Payment (APP) fraud, because the transaction is initiated and authorized by the victim, making detection and reimbursement more complex than in unauthorized fraud cases, such as those involving stolen credit card details.

Payment service providers (PSP), such as traditional banks or FinTech companies, operate a fraud detection model that flags suspicious transactions. These models are imperfect: some fraudulent transactions evade detection (false negatives), while some legitimate ones are incorrectly flagged (false positives). PSPs make an ex ante investment in their fraud detection model. A higher investment improves the model's accuracy, reduces the likelihood of false positives, and ultimately decreases the occurrence of fraud cases.

In our modeling, we distinguish between two types of fraud: Type A fraud, which is detectable, and Type B fraud, which is not detectable by PSP fraud models. For example, Type A fraud includes fraudulent transactions triggered by unusual patterns or inconsistencies, like invoice fraud, which is flagged by the system. In contrast, Type B fraud could involve romance scams, where the transaction appears legitimate, making it difficult for fraud detection models to identify.

The fraudster observes the average screening intensity across PSPs, which is influenced by both successful and unsuccessful fraud attempts. Based on this information, the fraudster adapts their strategy by increasing the share of Type B fraud, which is costlier, but

undetectable by PSP models. Consequently, as the average screening intensity increases, fraudsters shift towards more sophisticated, undetectable fraud tactics to avoid detection while still attempting to maximize their fraudulent returns. This dynamic creates a feedback loop that undermines the effectiveness of fraud detection efforts and influences the overall fraud landscape in the payment services market.

2.1 Timeline and Overview of the Fraud Screening Game

We now discuss our two-period fraud detection model. We start by providing an overview of the timing of the game. Figure 1 depicts the sequence of events. We discuss the details of the game after the timeline.

Time = 0

- *Consumers* choose a payment service.
- *Payment Service Providers* set a fee f for their payment service and compete à la Bertrand.
- *Payment Service Providers* set the screening intensity I_0 for their fraud model.

Time = 1

- *Payment Service Providers* screen transactions screen transactions, and any undetected fraud attempts result in realized fraud cases.
- *Fraudster* observes the average screening intensity of the PSPs $\bar{I}_0$ and sets the share of Type A fraud p_1 for period 2.

Time = 2

- *Payment Service Providers* screen transactions screen transactions, and any undetected fraud attempts result in realized fraud cases.

PSPs use the same fraud detection model in both period 1 and period 2. While we could allow the PSP to update its fraud model at $t = 1$, this would not lead to a different model because the PSP can only collect data about the fraudster's response in the subsequent period. For simplicity, we use a two-period model as it captures the main economic mechanism. An extension of this model could involve making it dynamic, allowing PSPs to decide each period, based on their model's performance, whether to update their fraud detection model. Key variables of the model, including those introduced here and in

subsequent sections, are tabulated in Table 1 for ease of reference. Appendix A contains all proofs.

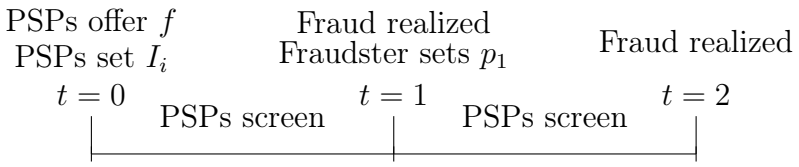


Figure 1
Timeline of events.

2.2 Details of the Fraud Screening Game

We consider a two-period economy with N identical payment service providers (PSPs), such as traditional banks or FinTech, who offer electronic payment services to consumers and no other financial products. All parties are risk neutral. We assume a zero discount rate throughout the analysis.

The economy also includes a single fraudster, who attempts to commit authorized push payment fraud on a share $q \in [0, 1]$ of all transactions. For each successful fraudulent transaction, the fraudster obtains revenue $V > 0$. There are two types of fraud: Type A and Type B. For example, the fraudster might impersonate legitimate entities to deceive consumers (e.g., via spoofing), or send fake invoices that appear legitimate.

Payment service providers could employ a screening technology to detect Type A fraud attempts. All type A fraud attempts which are labeled as fraud attempts are blocked by the PSP (true positives), however, the fraud attempts which are labeled as legitimate transactions (false negatives) result in actual fraud cases. The fraud success rate (likelihood of false negatives) ϕ_i is a function of the screening intensity I_i of PSP i.

$$\phi_i = Pr(FN|A)_i = 1^{\smile} I_i \tag{1}$$

The investment imposes a cost on PSPs that, for simplicity, we take to be quadratic: $(1/2)I^2$. Type B fraud is undetectable: $Pr(FN|B) = 1$.

2.3 Fraudster's choice of fraud type

We begin by analyzing the fraudster's choice of fraud type at $t = 1$. The fraudster determines at $t = 1$ the optimal share of type A fraud p_1, given the average fraud success rate in period $t = 0$ to maximize its profits:

$$\Pi_{fraudster} = (p_1(1\breve{}\bar{I}_0) + (1 - p_1))qV\breve{}(1/2)(1 - p_1)^2, \tag{2}$$

where $\bar{I}_0 = (1/N)\sum_{i=1}^{N} I_{i0}$ is the average screening intensity of the N PSPs in period $t = 0$ and V represents the fraud benefits. Switching to type B fraud imposes a cost on the fraudster that, for simplicity, we take to be quadratic: $(1/2)(1 - p_1)^2$. The fraudster chooses the share of type A fraud to maximize its profits 2. Taking the first order conditions with respect to p_1 yields and optimal share of type A fraud in period $t = 1$ of:

$$p_1^* = 1 - qV\bar{I}_0 \tag{3}$$

The fraudster choice depends on two main parameters: the fraud benefits V, the share of transactions exposed to fraud attempts, and the average screening intensity of the PSPs $\bar{I}_0$. Therefore, the fraud success rate in period 1 $\phi_i 1$ is the weighted average of the fraud success rate of type A and B:

$$\phi_{i1} = p_1 Pr(FN|A) + (1 - p_1)Pr(FN|B) = p_1(1 - I_{i0}) + (1 - p_1) = 1 - I_{i0} + qV\bar{I}_0 I_{i0}, \tag{4}$$

This equation shows that the fraud success rate increase in period $t = 1$ for higher average screening intensities because fraudsters dynamically respond to higher screening by increasing the share of undetectable type B fraud. We define the decline in fraud model effectiveness between $t = 0$ and $t = 1$ (measured as the change in the fraud success rate) as:

$$\Delta_{Model} = \phi_0 - \phi_1 \tag{5}$$

The model decay described above is closely related to the phenomenon of concept drift studied in the computer science and machine learning literature. Concept drift refers to the evolution of the underlying data-generating process in ways that invalidate a model's predictive accuracy [Bayram et al., 2022, Bolton and Hand, 2002]. Concept drift manifests through the growing prevalence of more sophisticated and less detectable fraud types—mirroring the increase in type B fraud in our model—which degrades model performance if left unaddressed. The result is a decline in predictive power akin to the model decay Δ_{Model} derived above.

In the next section, we examine how the foresight of model decay—that is, the PSPs' awareness that fraud models lose effectiveness over time—affects their screening intensities across different market structures.

2.4 PSP profits and screening intensities

In period $t = 0$, N identical PSPs set their screening intensity I_{i0} to maximize their profits:

$$\Pi_i = 2f - (\phi_{i0} + \phi_{i1})\,\gamma qV - (1/2)I_{i0}^2, \tag{6}$$

where $\phi_{i0} = (1 - I_{i0})$ and $\phi_{i1} = (1 - I_{i0})p_1^* + (1 - p_1^*)$. The first term on the left-hand side is the two-period fee revenue. γ is the PSP liability share of the total fraud costs. In this section we treat γ as an exogenous parameter. In the next section we examine the optimal γ choice by a financial regulator. The second term represents the expected fraud costs in period 1 and 2. The final term on the right-hand side is the screening cost to train the model in period $t = 0$. The first order condition of 30 with respect to I_{i0} is:

$$\frac{\partial \pi_i}{\partial I_0} = \gamma V \left(2 - \frac{2qV}{N} I_0 - \frac{qV}{N} \sum_{j \neq i} I_j\right) - I_0 = 0 \tag{7}$$

We assume that all payment service providers (PSPs) choose the same initial investment level, $I_{i0} = I$, implying symmetry among the PSPs. This leads to the average investment across all PSPs being $\bar{I} = I$. Substituting this assumption into the first-order conditions (FOC) yields:

$$I_0^* = \frac{2\gamma qVN}{N + \gamma q^2 V^2 (N+1)} \tag{8}$$

Finally, the N payment service providers (PSPs) compete à la Bertrand, resulting in zero profits for each PSP. The equilibrium payment service fee, denoted by f^*, is given by:

$$f^* = (1/2)\left(2 - 2I_0^* + qV(I_0^*)^2\right)\gamma qV - (I_0^*)^2 \tag{9}$$

Proposition 1. *In the symmetric equilibrium with N PSPs:*

1. *Each PSP chooses screening intensity $I_0^* = \frac{2\gamma qVN}{N+\gamma(qV)^2(N+1)}$, increasing in γV and N.*
2. *The equilibrium fee is $(2 - 2I_0^* + qV(I_0^*)^2)\,\gamma qV - (1/2)(I_0^*)^2$.*
3. *The decline in model effectiveness (measured as the increase in the fraud success rate) is given by $\Delta_{Model} = -qV(I_0^*)^2$.*

Proposition 1 shows that the optimal screening intensity I_0^* increase in the number of PSPs N. This is because individual PSPs do not internalize the model decay costs (Figure 2). Since I_0^* is increasing in N, the model decay Δ_{Model} is larger if N increases. [Hauswald and Marquez, 2006] study the effect of competition on IT investment in credit markets, and show that banks invest in screening because they try to 'steal business' from competitors. Since all banks follow identical strategies in equilibrium, such attempts are not successful and simply lead to overinvestment compared with a monopolist market structure. Our overinvestment results is not driven by business stealing. In contrast, PSPs overinvest because they do not internalize fully internalise the effect of future model decay.

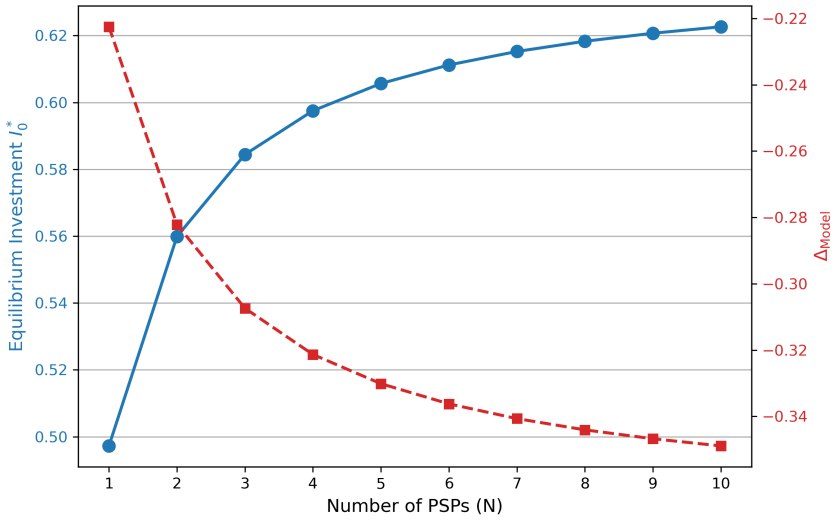


Figure 2
Competition, Screening Intensity, and Model Decay

Our research is related to the literature on IT investment in credit markets [Hauswald and Marquez, 2006, Vives and Ye, 2025]. Hauswald and Marquez [2006] show that banks invest in screening technology to 'steal business' from other banks, which could result in over-investment in screening in equilibrium. However, with more banks, the investment in screening technology for each bank falls because more competition reduces the returns to information acquisition. While the existing literature focuses on credit markets, our paper is the first to study investment in screening technologies in the payment markets.

3 Welfare, policy interventions, and consumer moral hazard

In this section we study the welfare implications of fraud screening. We first derive the social optimal screening level in the fraud game discussed in the previous sections. Next, we study the optimal liability rule a payment supervisor could set. Finally, we study the effect of consumer moral hazard - reduced vigilance due to high fraud reimbursement.
We make one simplifying assumption to keep the model simple and traceable:

$$\phi_{i1} = 1 - I_{i0} + \rho \bar{I_0}, \tag{10}$$

where $\rho \in [0, 1]$ is an exogenous model decay parameter. In this specification, the probability of fraud success is linearly increasing in the average PSP screening intensity. By contrast, under the more detailed probabilistic model in Equation (4), the fraud externality is nonlinear in both individual and average screening intensities, reflecting the endogenous fraudster response to aggregate screening. Despite these differences in functional form, both approaches capture the same core economic trade-off between individual screening effort and the model decay externalities. Crucially, the central economic intuition—that PSPs' efforts are interdependent through fraud detection spillovers—remains intact under either specification. As a result, the model's main results, including the characterization of optimal screening incentives and the comparative statics, are robust to the choice of functional form.

3.1 Social optimum screening investment

The welfare function for an economy with one PSP is:

$$\max_{\gamma} W = 2v - (\phi_0 + \phi_1) qV - \frac{1}{2}(I_0)^2, \tag{11}$$

where v is the utility consumers derive from using payment services, $\phi_0 = 1 - I_0$, $\phi_1 = 1 - I_0 + \rho I_0$. The second expression is the expected fraud costs, and the third expression the total fraud investment by the PSP.
We take the first order conditions of 11 with respect to I_0 to determine social optimum investment level:

$$I_W^* = qV\,(2 - \rho) \tag{12}$$

We compare the social optimum investment I^*_W with the optimum investment of a monopolistic PSP ($I^*_{W,N=1}$) and competitive PSPs ($I^*_{PSP,N>1}$).

Proposition 2. *The socially optimal screening investment level is given by* $I^*_W = qV\left(2 - \frac{\rho}{N}\right)$

1. $I^*_W > I^*_{PSP}$ *for* $\gamma < 1$ *and* $N = 1$
2. $I^*_W < I^*_{PSP,N>1}$ *for* $\gamma = 1$ *and* $N > 1$.

The first inequality reflects a *moral hazard* problem: when the liability parameter γ is less than one, a monopolistic PSP does not fully internalize the expected costs of fraud. Consequently, the PSP underinvests in screening relative to the social optimum. This underinvestment arises because the PSP bears only a fraction γ of the fraud losses, leading to insufficient incentives to allocate resources toward fraud prevention.

This mechanism closely parallels a central insight from the banking literature: financial intermediaries engage in privately costly, publicly unobservable screening and monitoring of borrowers. To ensure they exert efficient effort, it is necessary that they retain some exposure to the credit risk of the loans they originate [Pennacchi, 1988, Gorton and Pennacchi, 1995, Parlour and Plantin, 2008, Gryglewicz et al., 2024]. In both settings, partial liability or risk retention aligns incentives for privately informed agents to internalize the full social cost of adverse outcomes and thus invest efficiently in information acquisition and risk mitigation.

The second inequality illustrates an *overinvestment* phenomenon driven by a *model decay externality*. In a competitive market with multiple PSPs ($N > 1$) and full liability ($\gamma = 1$), each PSP is fully responsible for fraud losses but does not consider the negative externality its own screening imposes on others through model decay, captured by the parameter ρ. As each PSP intensifies screening to protect itself, the collective effect accelerates model decay, leading to an aggregate investment in screening that exceeds the social optimum. This overinvestment results from each PSP's failure to internalize the adverse impact of its actions on the shared fraud detection model's effectiveness.

A payment regulator might set a minimum investment that equals the social optimal I^*_W to increase welfare in the payment market. However, this policy might not me feasible if the screening investment is not feasible. We therefore consider a second policy option: setting a liability rule that sets the share of the fraud costs that the PSP has to reimburse the consumer.

3.2 Optimal fraud liability rule

The allocation of liability in financial fraud cases has become a policy lever in payment system regulation. Recent initiatives in the UK and EU attempted to increase the PSP liability in cases of authorized push payment fraud, with the aim of strengthening provider incentives to invest in screening technologies and protect consumers. By contrast, frameworks such as the U.S. Electronic Fund Transfer Act place comparatively less responsibility on financial institutions. These divergent approaches reflect ongoing uncertainty about how liability should be distributed across market participants to best mitigate fraud risks. In this section, we take the regulator's perspective and ask: what level of PSP liability maximizes overall welfare?

Formally, the liability share borne by PSPs is captured by the parameter $\gamma \in [0, 1]$. A higher γ means that a greater fraction of total fraud losses are internalized by PSPs, inducing stronger investment in screening. However, this comes at a cost: screening technologies are costly to implement, and excessive liability may lead to over-investment. The regulator's objective is to select γ to minimize the total cost of fraud to society, taking into account both direct fraud losses and screening expenditures. The optimal policy choice, γ^*, is the value that maximizes W.

Proposition 3. *The optimal investment fraction* $\gamma^* = \frac{2-\rho}{2N-\rho}$ *is:*

1. *Strictly decreasing in* N.
2. *Decreasing in the rate of model decay* ρ *for* $N \geq 1$.

We find that the regulator sets a lower γ^* in more competitive payment markets, since more competition increase the model decay externality. To prevent overinvestment in fraud technologies, the regulator sets a smaller $\gamma*$. An increase in the exogenous model decay parameter ρ diminishes the effectiveness of fraud models. A higher fraud decay parameter increases the fraud model screening externality. Therefore, the regulator sets a lower $\gamma*$ for higher values of ρ.

We also provide a numerical solutions to illustrate these findings. We therefore solve for γ^* numerically by minimizing $W(\gamma)$ over the unit interval. To implement this numerical solution, we calibrate the model using a baseline set of parameters reported in Table 1. We assume that consumers enjoy a utility of $v = 3$ from using payment services, a share of $q = 0.3$ consumers experiences a fraud attempt each period, and a fraud cost V of 3. We calculate welfare as a function of γ, using equation 8 to calculate the equilibrium value of fraud detection investment I_0^*. Figure 3 shows the welfare for different values of $\gamma \in [0, 1]$ and different market structures ($N = 1, 2, 5$).

The the optimal $\gamma = 1$ if there is one monopolist PSP. This is because the PSP internalizes the model decay and therefore that first best is to give the PSP the strongest incentives to screen by setting γ equal to 1. Increasing competition results in $\gamma^* < 0$, because competition makes PSPs overinvest in screening. The regulator faces a trade-off between providing the right incentives to PSPs to invest in fraud detection, and preventing overinvestment, which would reduce welfare. Therefore, regulators could better propose a partial liability rule in competitive payment markets.

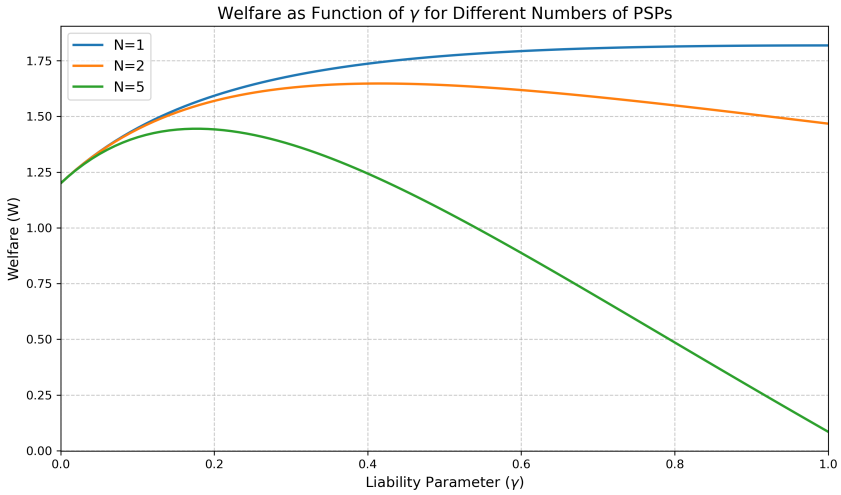


Figure 3
γ and welfare

3.2.1 Consumer moral hazard

One concern of shifting liability from consumers to PSPs is that consumers become less vigilant, which ultimately results in a higher share of transactions that are exposed to fraud attempts. In this section we will introduce consumer moral hazard to the model. We now introduce consumer moral hazard into the model. Consumers can remain vigilant at a private cost $c > 0$. Vigilance affects the probability of fraud attempts: the share of fraudulent transactions is q_l when consumers are vigilant and $q_h > q_l$ when they are not. This creates the following incentive compatibility constraint:

$$(q_h - q_l)(1-\gamma)(\phi_{i0} + \phi_{i1})V \geq c \tag{13}$$

where the left-hand side represents the net benefit of vigilance and the right-hand side captures the private cost. The constraint requires that consumers' expected benefit from vigilance outweighs their cost.

We maximize welfare W subject to the incentive compatibility constraint (ICC):

$$\max_{\gamma} W = 2v - (\phi_{i0} + \phi_{i1})q_l V - \frac{1}{2}N(I^*(\gamma))^2 \tag{14}$$

subject to:

$$(q_h - q_l)(1-\gamma)(\phi_{i0} + \phi_{i1})V \geq c, \tag{15}$$

where: $\phi_{i0} + \phi_{i1} = 2 - 2I^* + \rho I^*$ and $I^*(\gamma) = \gamma q V \left(1 + \frac{\rho}{N}\right)$.

Proposition 4 (Optimal Investment with Consumer Moral Hazard)**.** *The solution to the regulator's problem with consumer moral hazard yields distinct cases:*

1. ***When ICC does not bind*** *($c \leq \underline{c}$):*

$$\gamma^* = \frac{2-\rho}{2N-\rho} \tag{16}$$

where $\underline{c} \equiv (q_h - q_l)(1-\gamma^*)(2-(2-\rho)\gamma^* q V(1+\rho/N))V$.

2. ***When ICC binds*** *($c > \underline{c}$):*

$$\gamma_{ICC} = \frac{AB - 2B + \sqrt{(AB-2B)^2 + 4AB(2B-c)}}{2AB} \tag{17}$$

where $A \equiv (2-\rho)qV(2-\rho/N)$ *and* $B \equiv (q_h - q_l)V$.

The threshold $\underline{c}$ satisfies:

$$\underline{c} = (q_h - q_l)\left(1 - \frac{2-\rho}{2N-\rho}\right)\left(2 - \frac{(2-\rho)^2 qV}{2N-\rho}\left(1 + \frac{\rho}{N}\right)\right)V \tag{18}$$

The optimal liability rule in case 1 is strictly higher than the liability rule in case 2 ($\gamma^* > \gamma_{ICC}$)

The regulator's optimal liability policy reflects a trade-off between incentivizing the payment service provider (PSP) to screen for fraud and ensuring consumers remain diligent in protecting their credentials—a classic consumer moral hazard problem. When the cost of consumer effort c is sufficiently low (i.e., $c \leq \underline{c}$), the incentive compatibility constraint (ICC) does not bind, and the regulator can choose a liability level γ^* that optimally balances PSP screening incentives with fraud mitigation. This solution depends only on market parameters like the number of consumers N and the social value of transactions V. However, when $c > \underline{c}$, the ICC binds, meaning consumers are unwilling to exert sufficient care unless the PSP bears more liability. In this case, the regulator must raise γ to a level γ_{ICC} that restores consumer incentives, even though doing so may blunt the PSP's own screening effort. The threshold $\underline{c}$ therefore separates regimes where consumer behavior is self-disciplined from those where it must be subsidized via stricter provider liability.

Figure 4 illustrates this proposition with a numerical example. For a monopolist PSP ($N = 1$), the optimal γ is not 1 anymore. Also in more competitive markets, the optimal γ^* is lower than in an environment without consumer moral hazard. This creates a trade-off for the regulator: setting a γ value sufficiently high to give PSPs incentives to screen, but also setting γ not too high to prevent welfare-decreasing consumer moral hazard.

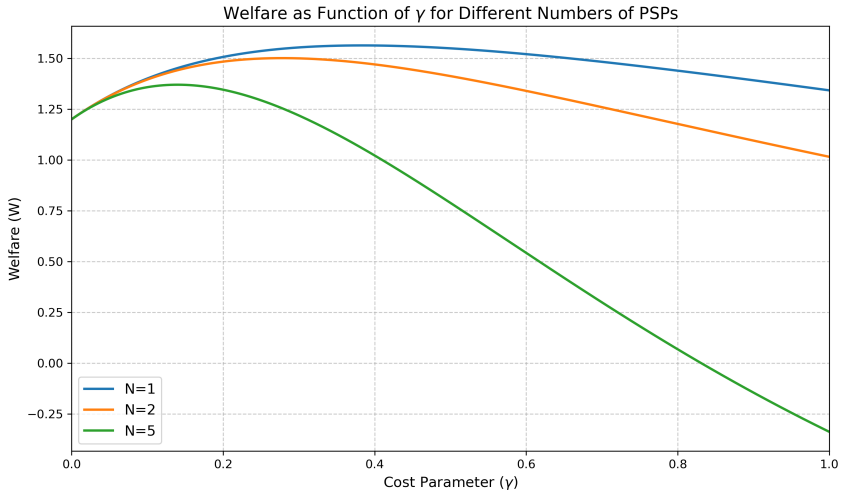


Figure 4
γ and welfare with consumer moral hazard

These findings caution payment regulators that with costly fraud screening and consumer moral hazard, fraud liability policies can lead to unintended consequences.

4 Fraud and account closures

In addition to fraud screening, PSPs can further reduce fraud risks by closing customer accounts following the detection of fraudulent activity. While the previous section's model applies to both the sending and receiving PSPs in a transaction, account closures primarily target customers on the receiving side, such as money mules. However, closures may also occur on the sending side if the PSP suspects the customer is complicit in the fraud (first-party fraud) or deems the customer too high-risk. Figure 5 illustrates the sequence of events under this extended model, incorporating account closures.

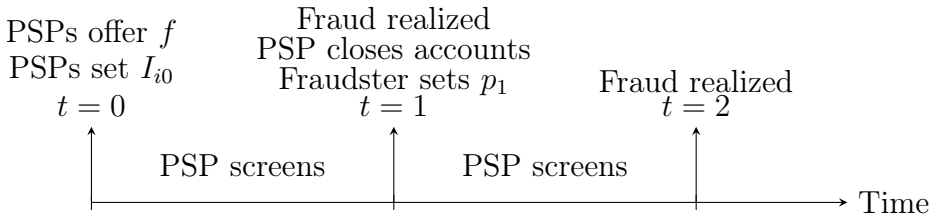


Figure 5
Timing with account closures

Consumers are heterogeneous in their exposure to fraud attempts. There are two types of consumers:

- High-risk consumers (share λ) are targeted with probability q_h.
- Low-risk consumers (share $1-\lambda$) are targeted with probability q_l

Thus, the unconditional probability of a fraud attempt is:

$$q_0 = \lambda q_h + (1-\lambda) q_l$$

Following a fraud event in period $t = 0$, PSPs may close accounts of consumers involved in successful frauds, should the expected returns from such accounts be negative, depending on the value of f and fraud costs V. Since high-risk consumers are more likely to be victims, they are also more likely to be removed, changing population composition by period $t = 1$: the population share of high-risk consumers evolves endogenously. Let $\bar{I}_0$ denote the average screening intensity in period $t = 0$. Then, the share of high-risk consumers who are not removed is proportional to:

$$\lambda \left[1 - q_h(1 - \bar{I}_0)\right],$$

while the surviving share of low-risk consumers is:

$$(1-\lambda) \left[1 - q_l(1 - \bar{I}_0)\right].$$

Thus, the overall fraud exposure in period $t = 1$ becomes:

$$q_1 = q_0 - (1 - I_{i0}) \left(\lambda q_h^2 + (1-\lambda) q_l^2\right) \tag{19}$$

Similarly, the surviving mass of consumers is:

$$m_1 = 1 - q_0(1 - I_{i0}) \tag{20}$$

And using Equation 10, we have:

$$\phi_{i1} = (1 - I_{i0}) + \rho \bar{I}_0 \tag{21}$$

In period $t = 0$, N identical PSPs set their screening intensity I_{i0} to maximize their profits:

$$\Pi_i = f(1 + m_1) - \gamma V(\phi_{i0} q_0 + \phi_{i1} q_1) - \frac{1}{2} I_{i0}^2 \tag{22}$$

The first-order condition for optimal screening investment is:

$$\frac{\partial \Pi_i}{\partial I_{i0}} = f\frac{\partial m_1}{\partial I_{i0}} - \gamma V\left[\frac{\partial \phi_{i0} q_0}{\partial I_{i0}} + \frac{\partial \phi_{i1} q_1}{\partial I_{i0}}\right] - I_{i0} \tag{23}$$

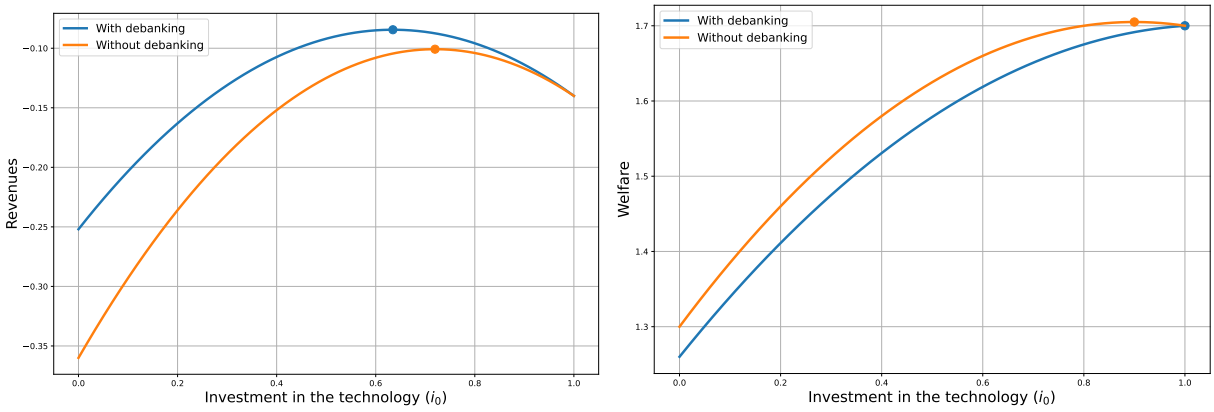
Proposition 5 (Optimal Investment with Debanking)**.** *With N=1, Equation 23 has a closed-form solution.*

$$I_{i0} = \frac{f q_0 - \gamma V(2-\rho)(\beta - q_0)}{2\beta V \gamma(\rho - 1) + 1} \tag{24}$$

where:

$$\beta = \lambda q_h^2 + (1-\lambda) q_l^2$$

Figure 6a plots the PSP's revenues and the optimal screening investment $I_0$when banks are permitted to debank consumers. Allowing account closures can either strengthen or weaken a PSP's incentive to invest in fraud detection, depending on the structure of service fees and liability rules. In particular, higher service fees make debanking more costly for PSPs, thereby encouraging greater upfront investment in fraud screening. Moreover, revenues are generally higher when debanking is allowed, indicating that banks may prefer removing high-risk consumers, an outcome that raises broader welfare concerns from the perspective of a social planner.



(a) Revenues **(b)** Welfare

Figure 6
Effects of Debanking

4.1 Increasing f: the case for cross-subsidies

In the presence of higher service fees f, account closures ("debanking") become more costly for PSPs. This is because closing an account means forfeiting future fee income from that customer. As a result, PSPs have stronger incentives to invest more in fraud detection upfront to avoid the need for costly debanking actions. This dynamic creates a welfare trade-off: Higher fees increase the cost of using payment services for consumers, but they also encourage PSPs to improve fraud detection, which reduces the likelihood of fraud events and the number of innocent consumers being debanked. Figure 7 illustrates this trade-off: as fees f increase, PSPs choose higher screening investments, resulting in better fraud prevention and a lower rate of account closures.

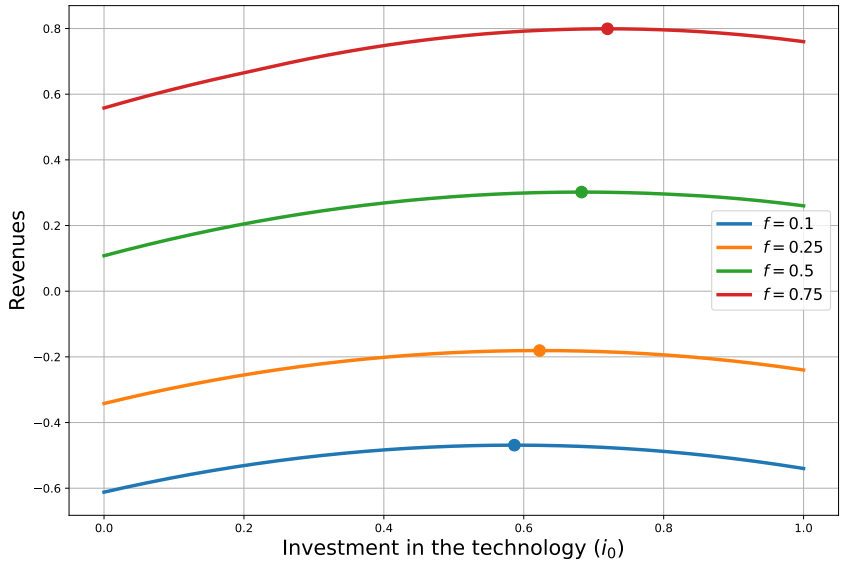
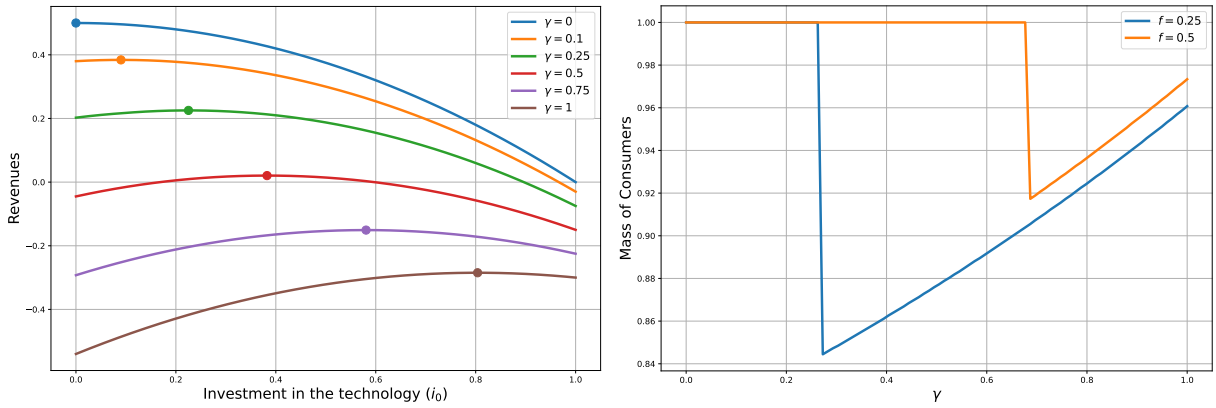


Figure 7
With Varying Fees

4.2 Debanking and γ: unexpected effects

The liability share γ, which governs how much fraud cost the PSP bears, also interacts with account closure dynamics in complex ways. Higher γ increases PSPs' incentive to screen more aggressively, as they are more financially responsible for fraud losses. However, when account closure is an option, PSPs may prefer to debank suspicious accounts rather than invest heavily in upfront screening.

Figures 8a and 8b highlight two important findings: As γ increases, PSPs' revenues decrease while their screening intensity increases. But higher γ reduces the mass of consumers, as more accounts are closed post-fraud events.



(a) Revenues **(b)** Debanking

Figure 8
γ's effect on debanking

Thus, increasing γ may unintentionally lead to more debanking, reducing financial access for consumers, especially those wrongly flagged as risky. In short, the effects of liability policies on welfare are ambiguous: stronger PSP liability improves screening incentives but can also magnify exclusion risks if PSPs respond primarily by debanking rather than by improving fraud detection models.

5 Liability between sending and receiving PSPs

In this section, we distinguish between sending and receiving PSPs, considering the distinct roles each plays in fraud prevention. The receiving PSP is the PSP that offers an account to the receiver of the fraudulent transaction. Fraudsters often recruit money mules or use false credentials to create payment account that are used to receive the money from fraudulent transactions. PSPs screen customers before offering them a payment account. Weak "Know Your Customer" screening procedures could therefore increase the number of accounts available for receiving fraudulent transactions and increase expected fraud losses. [5] The sending PSP screens transactions, as discussed in the previous sections and tries to detect fraudulent transactions. Since the sending and receiving PSP have distinct roles, payment service regulators have to decide whether to make the sending PSP, the receiving PSP or both liable for fraudulent transactions. In this section we therefore introduce γ_s and γ_r, where $0 \leq \gamma_s + \gamma_r \leq 1$. Setting $\gamma_r = 1$ provides strong incentives for the

[5] [on Banking Supervision, 2016] defines four essential elements necessary for a sound KYC programme: (i) customer acceptance policy; (ii) customer identification; (iii) ongoing monitoring of higher risk accounts; and (iv) risk management.

receiving PSP, while setting $\gamma_s = 1$ give strong incentives for the sending PSP to increase screening. From 7 October 2024, the UK Payment Systems Regulator (PSR) mandated sending and receiving PSPs to reimburse APP fraud victims 50:50.

While in practice many PSPs serve as both senders and receivers, we assume, for simplicity, an economy with one sending PSP and one receiving PSP. This simplification is justified by the empirical observation that most PSPs are structurally biased toward either sending or receiving roles. According to the UK Payment Systems Regulator [Payment Systems Regulator, 2024], Metro Bank had the highest value of APP scams sent per £ million of transactions in 2023 (£226), while Fintech firms like Wise and Revolut topped the rankings for APP scams received per £ million transactions—£974 and £756 respectively. This asymmetry suggests that PSPs often play a dominant role on one side of the transaction chain, validating our model's distinction between sending and receiving PSPs. In our model, we restrict to an economy in which the sending and receiving PSP operate in the same jurisdiction. Our analysis therefore does not apply to transactions to foreign banks or crypto accounts.

5.1 The receiving and sending PSP

We assume that a fraction $\lambda \in (0, 1)$ of customers at the receiving payment service provider (PSP) are bad (e.g., money mules), while the remaining $1-\lambda$ are good. To mitigate fraud risk, the receiving PSP invests in know-your-customer (KYC) screening with intensity $I_{receiving} \in [0, 1]$. We do not distinguish between ex ante KYC (e.g. Checking the identity of the account holder) and ex post KYC (e.g. on-going monitoring of accounts), in line with the Basel Committee KYC standards [on Banking Supervision, 2016]. Therefore, we assume that KYC screening has no impact on the total demand for payment services. A higher value of $I_{receiving}$ corresponds to more stringent KYC screening, reducing the likelihood that a bad customer could engage in fraudulent activities. Specifically, the probability that a bad customer engages in fraudulent activities is $1 - I_{receiving}$. The cost of screening is convex and given by $\frac{1}{2}I^2_{receiving}$. Under this setup, the expected share of bad customers who engage in fraudulent activities is $\lambda(1 - I_{receiving})$.
The profit function of the receiving PSP is:

$$\Pi_{receiving} = (2f - (\phi_0 + \phi_1)\gamma_r qV(\lambda \cdot (1 - I_{receiving})) - \frac{1}{2}I^2_{receiving}, \tag{25}$$

where $\phi_0 = 1 - I_{Sending}$, $\phi_1 = 1 - I_{Sending} + \rho I_{Sending}$.
The profit function of the sending PSP is:

$$\Pi_{sending} = (2f - (\phi_0 + \phi_1)\gamma_s qV(\lambda \cdot (1 - I_{receiving})) - \frac{1}{2} I^2_{sending}, \tag{26}$$

Proposition 6 (Equilibrium Screening Investments)**.** *In a payment system with sending and receiving PSPs, the equilibrium KYC investments are given by:*

1. *Receiving PSP's equilibrium investment:*

$$I^*_{receiving} = \frac{2A - (1-\rho)(2-\rho)AB}{1 - (1-\rho)(2-\rho)AB} \tag{27}$$

2. *Sending PSP's equilibrium investment:*

$$I^*_{sending} = (2-\rho) \cdot B \cdot \left(1 - I^*_{receiving}\right) \tag{28}$$

where:

$$A = \gamma_r qV\lambda,$$
$$B = \gamma_s qV\lambda.$$

The equilibrium screening investments characterized in the proposition reveal important economic intuitions about strategic interactions between payment service providers (PSPs). The receiving PSP's investment increases with the potential fraud losses, the effectiveness of its detection technology, and the prevalence of fraud , reflecting a standard cost-benefit calculation where higher stakes justify greater defensive spending. However, it decreases with the sending PSP's investment due to strategic substitutability - when senders do more screening, receivers can free-ride to some extent.

5.2 Welfare and the optimal liability rule

We now analyse how payment regulators could set γ_s and γ_r to maximise welfare:

$$\max_{\gamma_s, \gamma_r} W = 4v - (2(2)I^*_{sending})qV\lambda \cdot (1 - I^*_{receiving}) - \frac{1}{2}(I^*_{sending})^2 - \frac{1}{2}(I^*_{receiving})^2, \tag{29}$$

Since the welfare function is non-linear in γ_s and γ_r, we solve it numerically using the parameter values reported in Table 1. Figure 1 plots welfare as a function of γ_s, assuming $\gamma_r = 1 - \gamma_s$, for two values of the model decay parameter, $\rho = 0.3$ and $\rho = 0.5$. For this set of parameters, the welfare-maximizing allocation assigns a larger liability share

to the receiving PSP ($\gamma_s < \gamma_r$). However, the optimal allocation depends critically on the effectiveness of the fraud screening model. As ρ increases—reflecting greater model decay due to strategic adaptation by fraudsters—the regulator should shift more liability from the sending PSP to the receiving PSP.

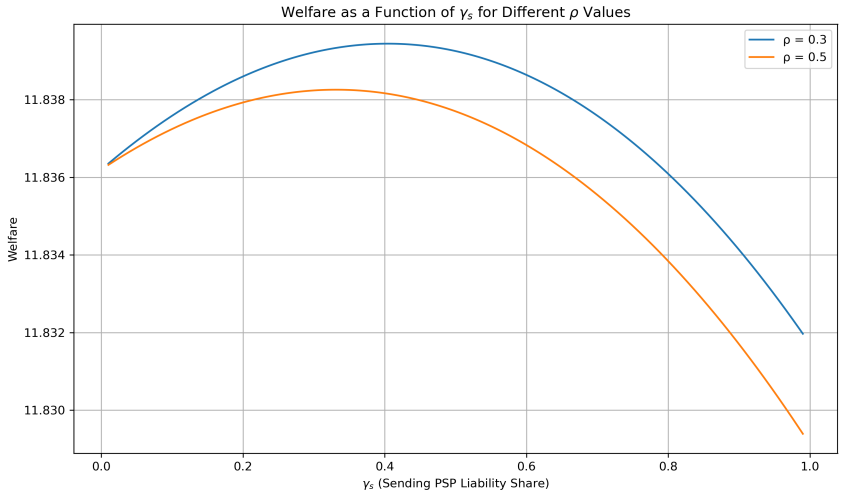


Figure 9
Welfare and γ_s

6 Conclusion

This paper develops a dynamic model of payment fraud in which fraudsters adapt to detection by evolving their tactics, causing a decline in model performance over time. We show that this dynamic distortion—akin to concept drift—undermines the long-run effectiveness of screening technologies. When PSPs do not bear the cost of fraud, screening is underprovided. While full liability can restore efficient screening under monopoly, it induces overinvestment in competitive markets, as each PSP's screening increases fraudster adaptation and imposes a negative externality on others.

Introducing consumer moral hazard further complicates the trade-offs in liability policy. Full reimbursement blunts consumer vigilance, suggesting a role for cost-sharing through partial reimbursement or consumer excess. We also show that account closures, which serve as an ex post screening mechanism, alter the population of users and feed back into ex ante fraud dynamics. While aggressive debanking can deter fraud, it may also reduce access to financial services. These institutional features underscore the importance of designing liability regimes that internalize both immediate and long-run behavioral responses from fraudsters, consumers, and PSPs.

Finally, we examine how liability should be shared between sending and receiving PSPs. Fraud prevention is a joint task: senders screen transactions, while receivers vet account holders. The optimal liability split depends on the relative decay in these two technologies. When fraudsters quickly adapt to transaction-level screening, placing more responsibility on receiving PSPs can improve welfare if know-your-customer checks degrade less rapidly in relative terms. This suggests that recent regulatory frameworks, such as the U.K.'s 50-50 split between sending and receiving PSPs, may not fully account for asymmetric decay in detection technologies. Future research could quantify adaptation dynamics more precisely and evaluate liability rules using field data from real-time payment networks.

References

Anup Agrawal, Jeffrey F Jaffe, and Jonathan M Karpoff. Management turnover and governance changes following the revelation of fraud. *The Journal of Law and Economics*, 42(S1):309–342, 1999.

Firas Bayram, Bestoun S Ahmed, and Andreas Kassler. From concept drift to model degradation: An overview on performance-aware drift detectors. *Knowledge-Based Systems*, 245:108632, 2022.

Richard J Bolton and David J Hand. Statistical fraud detection: A review. *Statistical science*, 17(3):235–255, 2002.

John H Boyd and Edward C Prescott. Financial intermediary-coalitions. *Journal of Economic theory*, 38(2):211–232, 1986.

Pushpita Chatterjee, Debashis Das, and Danda B Rawat. Digital twin for credit card fraud detection: Opportunities, challenges, and fraud detection advancements. *Future Generation Computer Systems*, 2024.

Stephen G Dimmock and William C Gerken. Predicting fraud by investment managers. *Journal of Financial Economics*, 105(1):153–173, 2012.

Stephen G Dimmock, William C Gerken, and Nathaniel P Graham. Is fraud contagious? coworker influence on misconduct by financial advisors. *The Journal of Finance*, 73(3): 1417–1450, 2018.

Alexander Dyck, Adair Morse, and Luigi Zingales. Who blows the whistle on corporate fraud? *The journal of finance*, 65(6):2213–2253, 2010.

Alexander Dyck, Adair Morse, and Luigi Zingales. How pervasive is corporate fraud? *Review of Accounting Studies*, 29(1):736–769, 2024.

Mark Egan, Gregor Matvos, and Amit Seru. The market for financial adviser misconduct. *Journal of Political Economy*, 127(1):233–295, 2019.

Gary B Gorton and George G Pennacchi. Banks and loan sales marketing nonmarketable assets. *Journal of monetary Economics*, 35(3):389–411, 1995.

Sebastian Gryglewicz, Simon Mayer, and Erwan Morellec. The dynamics of loan sales and lender incentives. *The Review of Financial Studies*, 37(8):2403–2460, 2024.

Robert Hauswald and Robert Marquez. Competition and strategic information acquisition in credit markets. *The Review of Financial Studies*, 19(3):967–1000, 2006.

Ludivia Hernandez Aros, Luisa Ximena Bustamante Molano, Fernando Gutierrez-Portela, John Johver Moreno Hernandez, and Mario Samuel Rodríguez Barrero. Financial fraud detection through the application of machine learning techniques: a literature review. *Humanities and Social Sciences Communications*, 11(1):1–22, 2024.

Vikramaditya Khanna, E Han Kim, and Yao Lu. Ceo connectedness and corporate fraud. *The Journal of Finance*, 70(3):1203–1252, 2015.

Basel Committee on Banking Supervision. Basel committee on banking supervision. *Guidelines Sound management of risks related to money laundering and financing of terrorismn (February 2016)*, 2016.

Christine A Parlour and Guillaume Plantin. Loan sales and relationship banking. *The Journal of Finance*, 63(3):1291–1314, 2008.

Christine A Parlour, Uday Rajan, and Haoxiang Zhu. When fintech competes for payment flows. *The Review of Financial Studies*, 35(11):4985–5024, 2022.

Payment Systems Regulator. Authorised push payment (app) scams performance report. Technical report, UK Payment Systems Regulator, July 2024. Available at: `https://www.psr.org.uk/`.

George G Pennacchi. Loan sales and the cost of bank capital. *The Journal of Finance*, 43(2):375–396, 1988.

Tomasz Piskorski, Amit Seru, and James Witkin. Asset quality misrepresentation by financial intermediaries: Evidence from the rmbs market. *The Journal of Finance*, 70 (6):2635–2678, 2015.

Paul Povel, Rajdeep Singh, and Andrew Winton. Booms, busts, and fraud. *The Review of Financial Studies*, 20(4):1219–1254, 2007.

Md Alamin Talukder, Majdi Khalid, and Md Ashraf Uddin. An integrated multistage ensemble machine learning model for fraudulent transaction detection. *Journal of Big Data*, 11(1):168, 2024.

Xavier Vives and Zhiqiang Ye. Information technology and lender competition. *Journal of Financial Economics*, 163:103957, 2025.

Tables

Table 1
Baseline Parameters for Numerical Analysis

Parameter	**Description**	**Range**	**Baseline Value**
v	Consumer utility from using payment services	[0, 10]	3
q	Share of transaction exposed to fraud	[0, 1]	0.3
V	Cost of a successful fraud attempt	[1, 10]	3
N	Number of competing PSPs	[1, 10]	1,2,5
ρ	Model decay parameter	[0, 1]	0.3,0.5
f	PSP payment service fee	[0, 10]	Endogenous
p_t	Share of Type A fraud in period t	[0, 1]	Endogenous
γ	PSP liability share (regulatory choice)	[0, 1]	Endogenous
I_{0i}	Fraud model screening intensity of PSP i, set at $t = 0$	[0, 1]	Endogenous
ϕ_{it}	Fraud success rate for PSP i in period t	[0, 1]	Endogenous

Appendix A: Proofs

A.1 Proof Proposition 1

Proposition 1. The equilibrium screening investment I_0^* is strictly increasing in N, the number of PSPs.

Proof. The equilibrium investment is given by:

$$I_0^* = \frac{2\gamma qVN}{N + \gamma(qV)^2(N+1)} \tag{B.1}$$

Let us define $A = 2\gamma qV$ and $B = \gamma(qV)^2$, both of which are strictly positive constants. Then the expression simplifies to:

$$I_0^* = \frac{AN}{N + B(N+1)} = \frac{AN}{N(1+B) + B}$$

We differentiate I_0^* with respect to N:

$$\frac{dI_0^*}{dN} = \frac{A[(1+B)N + B] - AN(1+B)}{[(1+B)N + B]^2}$$

Simplifying the numerator:

$$\begin{aligned} A[(1+B)N + B] - AN(1+B) &= A(1+B)N + AB - AN(1+B) \\ &= AB \end{aligned}$$

Therefore:

$$\frac{dI_0^*}{dN} = \frac{AB}{[(1+B)N + B]^2}$$

Since $A > 0$, $B > 0$, and the denominator is positive for all $N \geq 0$, it follows that $\frac{dI_0^*}{dN} > 0$. Hence, I_0^* is strictly increasing in N.

■

A.2 Proof Proposition 2

Proposition 2. $I_W^* > I_{PSP}^*$ for $\gamma < 1$ and $N = 1$.

We firstly derive the optimal PSP screening investment

PSP profit function is:

$$\Pi_i = 2f - (\phi_{i0} + \phi_{i1})\,\gamma qV - (1/2)I_{i0}^2, \tag{30}$$

where $\phi_{i0} = 1 - I_{i0}$, $\phi_{i1} = 1 - I_0 + \rho\bar{I}_0$ and $\bar{I}_0 = (1/N)\sum_{i=1}^{N} I_{i0}$.
We take the first order conditions with respect to I_{i0}:

$$\frac{d\Pi_i}{dI_{i0}} = \gamma qV\left(2 - \frac{\rho}{N}\right) - I_{i0} = 0$$

The optimum PSP investment level is:

$$I_{PSP}^* = \gamma qV\left(2 - \frac{\rho}{N}\right)$$

Now we could show that the optimal investment level of a monopolistic PSP ($N = 1$ is strictly less than the social optimum investment level if $\gamma < 1$:

$$I_W^* = qV\,(2 - \rho) > \gamma qV\,(2 - \rho) = I_{PSP,N=1}^*$$

Next, we compare the social optimum investment with the investment of competitive PSPs ($N > 1$) if $\gamma = 1$:

$$I_W^* = qV\,(2 - \rho) < qV\left(2 - \frac{\rho}{N}\right) = I_{PSP,N>1}^*$$

■

A.3 Proof Proposition 3

Proposition 3. The optimal investment fraction $\gamma^* = \frac{2-\rho}{2N-\rho}$ is: strictly decreasing in N and Decreasing in the rate of model decay ρ for $N \geq 1$.

We firstly substitute the optimal PSP investment into the welfare function:

$$\max_{\gamma} W = 2v - (\phi_{i0} + \phi_{i1})qV - \frac{1}{2}N(I^*(\gamma))^2, \tag{31}$$

where $\phi_0 = 1 - I_0^*$, $\phi_1 = 1 - I_0^* + \rho I_0^*$, and $I_{PSP}^* = \gamma qV\left(2 - \frac{\rho}{N}\right)$.
Next, we take the FOC of 31 with respect to γ and solve for γ^*:

$$\gamma^* = \frac{2 - \rho}{2N - \rho}$$

Part 2: Differentiating with respect to N yields:

$$\frac{\partial \gamma^*}{\partial N} = \frac{0 \cdot (2N - \rho) - (2 - \rho)(2)}{(2N - \rho)^2} = -\frac{2(2 - \rho)}{(2N - \rho)^2} < 0 \tag{32}$$

where the inequality follows because $2 - \rho > 0$ (since $\rho < 2$ by assumption) and the denominator is strictly positive.

Part 3: The cross partial derivative with respect to ρ is:

$$\frac{\partial \gamma^*}{\partial \rho} = \frac{(-1)(2N - \rho) - (2 - \rho)(-1)}{(2N - \rho)^2} = \frac{2(1 - N)}{(2N - \rho)^2} \tag{33}$$

For any meaningful network with $N \geq 1$, the numerator is non-positive while the denominator remains strictly positive, giving $\frac{\partial \gamma^*}{\partial \rho} \leq 0$. The inequality is strict when $N > 1$.

■

A.4 Proof Proposition 4

Proposition 4.

If ICC binds:
When the ICC binds, we solve the constraint with equality:

$$(q_h - q_l)(1 - \gamma_{ICC})(2 - 2I^* + \rho I^*)\gamma_{ICC} qV = c \tag{34}$$

Let $k \equiv qV\left(2 - \frac{\rho}{N}\right)$. Then:

$$(2 - 2I^* + \rho I^*) = 2 - (2 - \rho)I^* \tag{35}$$

$$= 2 - (2 - \rho)\gamma_{ICC} k \tag{36}$$

The constraint becomes:

$$(q_h - q_l)(1 - \gamma_{ICC})[2 - (2 - \rho)\gamma_{ICC} k]V = c \tag{37}$$

Step 2: Simplify and Rearrange

Let $B \equiv (q_h - q_l)V$ and $A \equiv (2 - \rho)k = (2 - \rho)qV\left(2 - \frac{\rho}{N}\right)$. Then:

$$B(1 - \gamma_{ICC})(2 - A\gamma_{ICC}) = c \tag{38}$$

Expanding yields the quadratic form:

$$-AB\gamma_{ICC}^2 + (2B - AB)\gamma_{ICC} + (2B - c) = 0 \tag{39}$$

Step 3: Solve Quadratic Equation

The solutions are:

$$\gamma_{ICC} = \frac{-(2B - AB) \pm \sqrt{(2B - AB)^2 + 4AB(2B - c)}}{-2AB} \tag{40}$$

Taking the economically meaningful root ($0 \leq \gamma_{ICC} \leq 1$):

$$\gamma_{ICC} = \frac{AB - 2B + \sqrt{(AB - 2B)^2 + 4AB(2B - c)}}{2AB} \tag{41}$$

Special Cases

1. **When** $c \to 0$:

$$\gamma_{ICC} \to \min\left(\frac{2}{A}, 1\right) = \min\left(\frac{2}{(2-\rho)qV(2-\rho/N)}, 1\right) \tag{42}$$

2. **When** $c \to c_{\max} = 2B$:

$$\gamma_{ICC} \to 0 \tag{43}$$

For case (1), when $c \leq \underline{c}$, the first-best solution satisfies the ICC. For case (2):

1. Substitute $I^* = \gamma qV(1 + \rho/N)$ into the binding ICC
2. Rearrange to obtain quadratic form $-AB\gamma^2 + (2B - AB)\gamma + (2B - c) = 0$
3. Solve for the economically meaningful root $\gamma \in [0, 1]$

The threshold $\underline{c}$ comes from evaluating the ICC at γ^*.

The constrained solution has the following properties:

- $\frac{\partial \gamma_{ICC}}{\partial c} < 0$ (higher costs reduce investment)
- $\frac{\partial \gamma_{ICC}}{\partial (q_h - q_l)} > 0$ (larger fraud differential increases investment)
- $\lim_{c \to \underline{c}^+} \gamma_{ICC} = \gamma^*$ (smoothly connects to unconstrained case)

■

A.5 Proof Proposition 5

Proposition 5.

In the case with only one PSP, we have the following:

$$\begin{aligned} m_1 &= 1 - q_0(1 - I_{i0}) \\ q_1 &= q_0 - \beta(1 - I_{i0}) \qquad \text{with } \beta = \lambda q_h^2 + (1-\lambda) q_l^2 \\ \phi_{i1} &= (1 - I_{i0}) + \rho I_{i0} \end{aligned}$$

Therefore:

$$q_1 \phi_{i1} = I_{i0}^2 \beta(\rho - 1) + I_{i0} \left(2\beta + \rho(q_0 - \beta) - q_0\right) + q_0 - \beta \tag{44}$$

And:

$$\frac{\partial q_1 \phi_{i1}}{\partial I_{i0}} = 2 I_{i0} \beta(\rho - 1) + 2\beta + \rho(q_0 - \beta) - q_0 \tag{45}$$

As a result:

$$\frac{\partial \Pi_i}{\partial I_{i0}} = f \frac{\partial m_1}{\partial I_{i0}} - \gamma q_0 V \left[\frac{\partial \phi_{i0} q_0}{\partial I_{i0}} + \frac{\partial \phi_{i1} q_1}{\partial I_{i0}} \right] - I_{i0} \tag{46}$$

$$= -I_{i0} \left[2\beta V \gamma(\rho - 1) + 1\right] + f q_0 - \gamma V (2 - \rho)(\beta - q_0) \tag{47}$$

Finally,

$$I_{i0} = \frac{f q_0 - \gamma V (2 - \rho)(\beta - q_0)}{2\beta V \gamma(\rho - 1) + 1} \tag{48}$$

■

A.6 Proof Proposition 6

Proposition 6.

The receiving PSP chooses $I_{\text{receiving}}$ to maximize:

$$\Pi_{\text{receiving}} = 2f - (\phi_0 + \phi_1)\gamma_r q V \lambda (1 - I_{\text{receiving}}) - \frac{1}{2} I_{\text{receiving}}^2,$$

which yields the first-order condition:

$$I_{\text{receiving}} = (\phi_0 + \phi_1) \cdot \gamma_r q V \lambda.$$

Substituting $\phi_0 + \phi_1 = 2 - (1 - \rho)I_{\text{sending}}$, we obtain:

$$I_{\text{receiving}} = [2 - (1 - \rho)I_{\text{sending}}] \cdot A.$$

The sending PSP chooses I_{sending} to maximize:

$$\Pi_{\text{sending}} = 2f - (\phi_0 + \phi_1)\gamma_s qV\lambda(1 - I_{\text{receiving}}) - \frac{1}{2}I^2_{\text{sending}},$$

with first-order condition:

$$I_{\text{sending}} = (2 - \rho) \cdot \gamma_s qV\lambda(1 - I_{\text{receiving}}) = (2 - \rho) \cdot B \cdot (1 - I_{\text{receiving}}).$$

Substituting this expression for I_{sending} into the equation for $I_{\text{receiving}}$ and simplifying yields equation (27). Substituting the solution for $I^*_{\text{receiving}}$ back into the FOC for I_{sending} gives equation (28).

■