When Carry Crowds: Endogenous Foreign Exchange Crash Risk and Capital Inflows

Lei Pan[*,†]

Abstract

We develop a tractable model of foreign portfolio inflows into emerging-market local-currency assets when exchange-rate crash risk is endogenous to positioning. A representative investor funds a carry trade at a global rate and chooses the inflow position under mean–variance preferences. The exchange rate features normal shocks and a crash jump whose probability rises with inflows. Under a rare-crash approximation, the investor's objective is quartic in the position and the first-order condition is cubic, delivering closed-form characterization and conditions for uniqueness. When carry amplification dominates locally but fragility dominates globally, non-concavity generates multiple stationary points and sudden-stop–type regime switching. We study the policy instrument—a wedge on foreigners' after-cost return—and show how they reduce inflows, dampen tail risk, and can eliminate multiplicity.

Keywords: Global financial cycle; emerging markets; endogenous crash risk; tail risk; macroprudential policy

JEL Classification: D81; F31; F32; G11; G15

[*]School of Accounting, Economics and Finance, Curtin University, Australia. Email: lei.pan@curtin.edu.au
[†]Centre for Development Economics and Sustainability, Monash University, Australia.

1 Introduction

In the upswing of a global financial cycle, the trade that channels money into emerging markets (EMs) can look almost dull. A leveraged investor borrows cheaply in a major funding currency, buys a local-currency bond in an EM, and counts on the carry—the interest differential—plus a stable or appreciating exchange rate. For long stretches, nothing dramatic happens. Then, seemingly all at once, the same positions are cut, the exchange rate jumps, and what looked like a steady yield pick-up is revealed to be a strategy that earns small gains while occasionally losing big. This pattern—procyclical inflows in good global times and abrupt reversals when global conditions tighten—is by now a familiar description of portfolio flows and exchange-rate risk in EMs (Rey 2015, Miranda-Agrippino and Rey 2020, Forbes and Warnock 2012, Fratzscher 2012).

A central feature of this environment is that the "global financial cycle" is not just about the level of world interest rates. It is also about intermediaries' risk appetite, balance-sheet capacity, and the price of bearing risk (Bruno and Shin 2015, Adrian and Shin 2010). When global funding is easy and measured risk is low, cross-border balance sheets expand and gross capital flows surge; when funding tightens or risk appetite evaporates, the same balance sheets shrink and flows retrench (Broner et al. 2013, Forbes and Warnock 2012). Empirically, common "push" forces—global liquidity and global risk—explain a large share of the high-frequency variation in portfolio flows (Fratzscher 2012, Ahmed and Zlate 2014). In that sense, the question for an EM is rarely whether it will be exposed to global shocks, but rather how that exposure maps into domestic asset prices, currency risk, and (crucially) the probability of disruptive tail events.

This paper offers a deliberately stripped-down theoretical framework for thinking about that mapping. The model is built for transparency: it keeps the basic carry-trade arithmetic but adds one element that is often discussed informally and modeled less often in closed form—that tail risk is endogenous. The key mechanism is simple. When foreign portfolio inflows become large, they can compress risk premia and support the currency in the short run, yet simultaneously increase fragility (for example, by encouraging leverage, maturity mismatch, crowded positioning, or procyclical risk-taking). As a result, the probability of a sharp exchange-rate crash can rise with the scale of inflows. In short, the same force that makes the trade look attractive in tranquil periods can also plant the seeds of a crash when conditions reverse.

To formalize this idea, we consider a two-period environment in which a representative foreign investor chooses an inflow position in a local-currency EM security. The investor finances the position at a global funding rate and faces mean–variance preferences in the spirit of the classical portfolio problem (Markowitz 1952). The novelty is in the exchange-rate process. Specifically, the gross exchange-rate change has a normal component and a crash jump, and the crash probability is an increasing function of the inflow position. This yields two sharp implications. First, because crash risk rises convexly with inflows, the investor's objective becomes a quartic polynomial in position size and the first-order condition is cubic. Second, the model naturally delivers regime-like behavior: for parameter values in which "carry amplification" dominates locally but fragility dominates globally, the objective can become non-concave, opening the door to multiple candidate inflow levels and discontinuous switches between high-inflow and low-inflow outcomes. This is a parsimonious way to generate "sudden stops" as a change in the location of the global maximizer rather than as an imposed occasionally binding constraint (Calvo 1998, Bianchi 2011).

The setup speaks to several strands of the international finance literature. A first strand documents the global financial cycle and its consequences for monetary autonomy and cross-border leverage (Rey 2015, Miranda-Agrippino and Rey 2020, Bruno and Shin 2015). A second strand studies the macroeconomics of capital-flow bonanzas, stops, and crises (Calvo 1998, Forbes and Warnock 2012, Gourinchas and Obstfeld 2012, Broner et al. 2013). A third strand focuses on currency carry trades and the empirical fact that carry returns are exposed to crash risk and negative skewness (Brunnermeier et al. 2009, Menkhoff et al. 2012, Jurek 2014, Burnside et al. 2011, Lustig et al. 2011). Finally, a growing policy literature asks when and how macroprudential tools or capital flow management measures can reduce the buildup of systemic risk (Ostry et al. 2010, 2012, Korinek 2011, Jeanne and Korinek 2010, Farhi and Werning 2014, 2016, Bianchi and Mendoza 2018, Magud et al. 2018).

The model is closest in spirit to theories that emphasize the interaction between leveraged balance sheets and asset prices. Funding conditions affect risk-taking, and risk-taking feeds back into prices and risk premia (Adrian and Shin 2010, Brunnermeier and Pedersen 2009, Geanakoplos 2010). In exchange-rate markets, limited risk-bearing capacity implies that capital flows can move exchange rates and risk premia. The present framework takes these insights and embeds them in a minimal carry-trade environment with one additional object: an explicit, position-dependent crash probability. This is also consistent with disaster-risk approaches that explain high premia with a small probability of large adverse outcomes (Barro 2006, Gourio 2012). Here, however, the disaster probability is not merely a state variable: it is partly a function of the position chosen in equilibrium.

Three empirical regularities motivate the modeling choices. First, gross inflows to EMs are large, volatile, and strongly procyclical, with sharp retrenchments during crises (Broner et al. 2013, Forbes and Warnock 2012). Second, global "push" forces—notably global risk and liquidity—are central for explaining the time variation in flows (Fratzscher 2012, Ahmed and Zlate 2014, Miranda-Agrippino and Rey 2020). Third, currency carry strategies earn premia that are difficult to reconcile with uncovered interest parity, and a prominent explanation is compensation for crash risk or tail events (Brunnermeier et al. 2009, Menkhoff et al. 2012, Jurek

2014, Farhi et al. 2009). The literature has emphasized different facets of this last point: some argue that large negative payoffs occur in bad times when marginal utility is high (Lustig et al. 2011); others emphasize that hedging crash events can explain a nontrivial share of carry premia (Jurek 2014); and still others debate the extent to which peso events are best interpreted as rare disasters versus rare states with high stochastic discount factors (Burnside et al. 2011). Rather than take a stand on the full asset-pricing decomposition, this paper isolates one mechanism that is plausible in many interpretations. In particular, crash risk rises when positioning becomes crowded.

On the policy side, there is a clear tension. On one hand, open capital accounts can support financial development and allow risk sharing; on the other, inflow booms can amplify leverage cycles and increase the likelihood of costly crises (Gourinchas and Obstfeld 2012). A large body of work rationalizes prudential capital controls or macroprudential regulation as Pigouvian tools that internalize crisis externalities (Jeanne and Korinek 2010, Korinek 2011, Bianchi 2011, Bianchi and Mendoza 2018). From a practical perspective, policy discussions often distinguish between broad capital controls, targeted inflow taxes, and prudential tools that operate through margins, reserve requirements, or leverage constraints (Ostry et al. 2010, 2012, Magud et al. 2018). The model here contributes to this debate by offering a simple environment where a reduced-form wedge on foreigners' after-cost return affects not only the level of inflows but also the slope of inflows with respect to global conditions and the implied tail probability.

This paper makes three important contributions. First, it introduces a tractable mechanism that links global-cycle-driven inflows to endogenous foreign exchange (FX) crash tail risk. Crash probability is increasing in the inflow position, so tail risk is not an exogenous "state of the world" but a function of equilibrium portfolio choice. This delivers a clean elasticity of tail probability with respect to inflows: when the endogenous component dominates baseline risk, tail probability scales approximately with the square of inflows, and the associated log–log elasticity approaches two.

Second, the paper provides a fully explicit characterization of the investor's problem and equilibrium inflows. With mean–variance preferences and rare-crash approximation, the objective becomes a quartic polynomial in the inflow position and the first-order condition is cubic. This yields closed-form expressions for candidate inflow levels and transparent conditions under which the solution is unique. When those conditions fail, the same closed-form structure clarifies how non-concavity can produce multiple candidate inflow "regimes" and discontinuous switches between them, offering a simple analytical notion of sudden stops without imposing an occasionally binding collateral constraint.

Third, the present paper clarifies how macroprudential policy can shape both the level of inflows and the buildup of tail risk. A reduced-form wedge that lowers foreigners' after-cost return (interpretable as a tax, reserve requirement, or regulatory wedge) reduces equilibrium inflows and therefore reduces endogenous crash probability. Moreover, by damping the carry-amplification channel, the wedge can restore global concavity and eliminate multiplicity, thereby removing a specific source of discontinuous regime switching. An equivalent conclusion holds for a VaR-style margin constraint: tighter margins cap the position size precisely in the region where endogenous jump risk makes the variance steep in inflows.

Section 2 presents the setup and the exchange-rate process with endogenous crash probability. More specifically, we first derives the investor's objective, proves existence and uniqueness results, and provides the closed-form characterization of equilibrium inflows. Then, we discusses the risk-premium analogue, tail risk elasticity, and the amplification mechanism. Finally, we analyzes the effects of macroprudential wedges and margin-style tools and records comparative statics with respect to the global financial cycle. Section 3 concludes.

2 Model

2.1 Set up

Time is $t \in \{0, 1\}$. There is a representative foreign investor (the marginal buyer) and an emerging-market (EM) local-currency security. Throughout, variables are expressed in gross form.

Let $e_t > 0$ denote the spot exchange rate defined as foreign currency per one unit of EM currency. Hence an increase in e is an EM currency appreciation from the foreign investor's point of view.

A scalar global state $g \in (0, 1]$ summarizes the global financial cycle. Higher g means easier global conditions (lower funding frictions and higher effective risk tolerance). The global state affects: i) the foreign funding gross rate $R^*(g) > 1$; ii) the investor's mean–variance risk-aversion coefficient $\lambda(g) > 0$; iii) the baseline volatility of normal-times currency fluctuations $\sigma_\varepsilon(g) > 0$; and iv) the baseline crash probability $\bar{\pi}(g) \in (0, 1)$. To keep the comparative statics transparent, we impose sign restrictions later.

At $t = 0$ the foreign investor allocates an amount $F \geq 0$ (in foreign currency units) into a local-currency security with gross local-currency return $R > 1$ between $t = 0$ and $t = 1$. This R can be interpreted as the gross return on a one-period EM local-currency bond or a broad EM LC security index.[1]

[1] Keeping R as a primitive matches the logic: the core object is the foreign-currency excess payoff that combines the local return and the exchange-rate change. One may endogenize R or the bond price; Section **??** sketches a pricing extension.

The foreign investor funds F at the global gross rate $R^*(g)$. A host-country macroprudential wedge $\tau \in [0, 1)$ acts like a tax on the local return component received by foreigners.[2]

The foreign-currency net excess payoff (relative to funding) from the EM investment is:

$$K \equiv F\left[(1-\tau)\frac{e_1}{e_0}R - R^*(g)\right] - \frac{\kappa}{2}F^2. \tag{1}$$

where e_1/e_0 is the gross currency appreciation factor (foreign per EM); $\kappa > 0$ is a convex "balance-sheet" or price-impact cost (captures limited depth, intermediation cost, or internal leverage penalties).

2.2 Normal times vs. crash times, and endogenous fragility

The key is that the exchange-rate change contains both a normal innovation and a crash jump whose probability depends on inflows. Specifically,

$$\frac{e_1}{e_0} = \underbrace{1 + \mu_e(g) + \chi F + \varepsilon}_{\text{normal component}} - \underbrace{J\,1_{\{C=1\}}}_{\text{crash jump}}, \tag{2}$$

where $\mu_e(g) \in R$ is baseline expected appreciation (higher means expected EM appreciation); $\chi \geq 0$ captures a portfolio-balance channel: larger inflows tilt demand toward EM currency and raise expected appreciation (imperfect sterilization, slow-moving arbitrage, etc.); ε is a normal-times shock with $\mathbb{E}[\varepsilon] = 0$ and $V(\varepsilon) = \sigma_\varepsilon^2(g)$; and $C \in \{0, 1\}$ is a crash indicator. If $C = 1$, the EM currency suffers a discrete depreciation of size $J > 0$ in gross terms.[3]

The crash probability is endogenous:

$$\mathbb{P}(C = 1 \mid F, g) = \pi(F, g) \equiv \bar{\pi}(g) + \eta F^2, \tag{3}$$

with $\eta > 0$ measuring fragility: larger inflows raise the likelihood of an FX crash (think currency mismatch, fragile funding, crowded carry, or procyclical dealer balance sheets). The quadratic form is chosen for tractability; it is the simplest smooth specification that generates strong non-linearities.

Assumption 1 (Feasible crash probability)**.** *There exists $\bar{F} > 0$ such that the equilibrium inflow satisfies $F \in [0, \bar{F}]$ and $\pi(F, g) \in (0, 1)$ for all (F, g) in this set.*

Remark 1. Assumption 1 is mild. Since $\pi(F, g) = \bar{\pi}(g) + \eta F^2$, one can pick $\bar{F}$ such that $\eta \bar{F}^2 < 1 - \sup_g \bar{\pi}(g)$.

[2] This wedge is deliberately reduced-form. It can capture a tax on foreign holdings, an unremunerated reserve requirement, or a regulatory wedge that lowers the effective after-cost return to foreign intermediaries.

[3] Because e is foreign per EM, a crash means e_1/e_0 falls; hence the subtraction in Equation (2).

2.3 Investor problem: mean–variance with endogenous tail risk

The foreign investor is mean–variance. Let $\theta > 0$ be a taste shifter on expected payoff (a "mean weight") and let $\lambda(g) > 0$ be risk aversion. The investor chooses $F \geq 0$ to maximize:

$$U(F; g, \tau) \; \equiv \; \theta\, \mathbb{E}[K \mid F, g] \; - \; \frac{\lambda(g)}{2} V(K \mid F, g). \tag{4}$$

where θ measures how aggressively the investor trades off mean vs. risk (higher θ increases desired exposure); and $\lambda(g)$ is risk aversion (higher λ reduces desired exposure); it moves with global conditions.

Define the random per-unit excess return in foreign currency:

$$X \; \equiv \; (1-\tau)\frac{e_1}{e_0} R \; - \; R^*(g). \tag{5}$$

Then $K = FX - \frac{\kappa}{2}F^2$. Conditional on (F, g),

$$\mathbb{E}[K] = F\, \mu_X(F, g, \tau) \; - \; \frac{\kappa}{2}F^2, \qquad V(K) = F^2\, \sigma_X^2(F, g, \tau), \tag{6}$$

where $\mu_X = \mathbb{E}[X]$ and $\sigma_X^2 = V(X)$.

Using Equations (2)–(3), and the independence of ε and C, we compute:

$$\mu_X(F, g, \tau) = (1-\tau)R\,(1 + \mu_e(g) + \chi F - J\pi(F, g)) \; - \; R^*(g), \tag{7}$$

$$\sigma_X^2(F, g, \tau) = (1-\tau)^2 R^2 \left(\sigma_\varepsilon^2(g) \; + \; J^2\, \pi(F, g)\big(1 - \pi(F, g)\big)\right). \tag{8}$$

The term $J\pi(F, g)$ in Equation (7) is a crash-expected-loss adjustment: higher crash probability lowers expected carry. The term $J^2\pi(1-\pi)$ in Equation (8) is the jump contribution to variance.

For closed-form characterization, it is convenient to work with the standard "rare crash" approximation:

Assumption 2 (Rare crash approximation)**.** *In the relevant equilibrium region, $\pi(F, g)$ is small enough that $1 - \pi(F, g) \approx 1$ in the second-order term of Equation (8). Hence,*

$$\sigma_X^2(F, g, \tau) \; \approx \; (1-\tau)^2 R^2 \left(\sigma_\varepsilon^2(g) + J^2 \pi(F, g)\right). \tag{9}$$

Remark 2. Assumption 2 is an approximation that preserves the qualitative economics while yielding a polynomial objective. One can drop it and keep $\pi(1-\pi)$; the proofs go through with heavier algebra (the degree of the polynomial increases).

2.4 Reduced-form quartic objective and a cubic first-order condition

Substituting (6) into Equation (4) yields

$$: U(F; g, \tau) = \theta\left(F\mu_X(F, g, \tau) - \frac{\kappa}{2}F^2\right) - \frac{\lambda(g)}{2}F^2\sigma_X^2(F, g, \tau). \tag{10}$$

Under Assumption 2 and $\pi(F, g) = \bar{\pi}(g) + \eta F^2$, define the following convenient coefficients:

$$A(g, \tau) \equiv (1 - \tau)R\,(1 + \mu_e(g) - J\bar{\pi}(g)) - R^*(g), \tag{11}$$

$$B(\tau) \equiv (1 - \tau)R\,\chi, \tag{12}$$

$$C(g, \tau) \equiv (1 - \tau)R\,J\eta, \tag{13}$$

$$S(g, \tau) \equiv (1 - \tau)^2R^2\left(\sigma_\varepsilon^2(g) + J^2\bar{\pi}(g)\right), \tag{14}$$

$$D(\tau) \equiv (1 - \tau)^2R^2\,J^2\eta. \tag{15}$$

where A is baseline expected excess carry (net of baseline crash losses); B is the strength of inflow-driven expected appreciation; C is the marginal expected crash loss induced by inflows; S is baseline variance; and D is how strongly inflows raise variance via endogenous crash probability.

Using Equations (11)–(15), the mean and variance simplify to:

$$\mu_X(F, g, \tau) = A(g, \tau) + B(\tau)F - C(g, \tau)F^2, \qquad \sigma_X^2(F, g, \tau) = S(g, \tau) + D(\tau)F^2. \tag{16}$$

Substitute (16) into Equation (10) to obtain a quartic polynomial in F:[4]

$$U(F; g, \tau) = \theta\left(AF + BF^2 - CF^3 - \frac{\kappa}{2}F^2\right) - \frac{\lambda(g)}{2}F^2\left(S + DF^2\right). \tag{17}$$

Differentiating Equation (17) yields the first-order condition:

$$U_F(F; g, \tau) = \theta\left(A + (2B - \kappa)F - 3CF^2\right) - \lambda(g)\left(SF + 2DF^3\right) = 0. \tag{18}$$

This is a cubic equation in F.

[4] Here and below, (g, τ) arguments of A, C, S are omitted when unambiguous.

2.5 Existence, uniqueness, and closed-form characterization

Lemma 1 (Concavity for large F and existence of an optimum)**.** *Suppose $D(\tau) > 0$ and $\lambda(g) > 0$. Then $\lim_{F\to\infty} U(F; g, \tau) = -\infty$. Moreover, if $A(g, \tau) > 0$, there exists at least one optimal inflow level $F^\star(g, \tau) \in (0, \infty)$ satisfying Equation (47).*

Proof. Part 1: Limit behavior (coercivity): We examine the asymptotic behavior of $U(F)$ as $F \to \infty$. Let us expand the terms in Equation (1) and group them by powers of F. The term with the highest degree determines the limit.

Expanding the expression:

$$
\begin{aligned}
U(F) &= \theta AF + \theta\left(B - \frac{\kappa}{2}\right)F^2 - \theta CF^3 - \frac{\lambda S}{2}F^2 - \frac{\lambda D}{2}F^4 \\
&= -\frac{\lambda D}{2}F^4 - \theta CF^3 + \left[\theta\left(B - \frac{\kappa}{2}\right) - \frac{\lambda S}{2}\right]F^2 + \theta AF
\end{aligned}
$$

The polynomial is of degree 4. Since $\lambda > 0$ and $D > 0$, the coefficient of the leading term F^4 is strictly negative:

$$-\frac{\lambda D}{2} < 0$$

Therefore, the tail behavior is dominated by this negative quartic term:

$$\lim_{F\to\infty} U(F) = -\infty$$

This proves the first part of the lemma. Intuitively, as the position size F grows, the variance of the position (which scales with F^4 due to the endogenous crash probability) grows faster than the expected return (which scales at most with F^2 or F^3). The risk penalty eventually overwhelms any expected gain.

Part 2: Existence of an interior optimum: We seek to maximize $U(F)$ on the domain $F \in [0, \infty)$.

1. **Continuity:** $U(F)$ is a polynomial, so it is continuous and differentiable everywhere.
2. **Boundary at infinity:** From Part 1, we know $\lim_{F\to\infty} U(F) = -\infty$. This implies that there exists some large number $\bar{F}$ such that for all $F > \bar{F}$, $U(F) < 0$.
3. **Value at zero:** Evaluating the objective at $F = 0$:

$$U(0) = 0$$

4. **Behavior near zero:** We analyze the derivative $U'(F)$ at $F = 0$. Differentiating $U(F)$:

$$U'(F) = \theta(A + (2B - \kappa)F - 3CF^2) - \lambda(SF + 2DF^3)$$

Evaluating at $F = 0$:

$$U'(0) = \theta(A + 0 - 0) - \lambda(0) = \theta A$$

Given the assumption $\theta > 0$ and the condition $A > 0$, we have:

$$U'(0) > 0$$

Conclusion: Since $U(0) = 0$ and $U'(0) > 0$, the utility function is strictly increasing at the origin. Thus, there exists some small $\epsilon > 0$ such that $U(\epsilon) > U(0) = 0$.

Since $U(F)$ increases initially but eventually approaches $-\infty$, it must achieve a global maximum somewhere in the interval $(0, \infty)$. The maximum cannot be at $F = 0$ (because $U(\epsilon) > U(0)$) and cannot be at infinity.

Therefore, there exists an optimal finite investment level $F^* \in (0, \infty)$. Being an interior optimum of a differentiable function, F^* must satisfy the first-order condition $U'(F^*) = 0$, which corresponds to Equation (47)

Note on concavity: The lemma mentions "Concavity for large F." We can verify this by taking the second derivative:

$$U''(F) = \theta(2B - \kappa - 6CF) - \lambda(S + 6DF^2)$$

As $F \to \infty$, the term $-6\lambda DF^2$ dominates. Since $\lambda, D > 0$, $U''(F) \to -\infty$, ensuring the function is concave for sufficiently large F. □

2.5.1 Uniqueness under a parameter restriction

The cubic equation (47) can in general admit multiple positive roots. A simple sufficient condition for uniqueness is that U is strictly concave on $[0, \infty)$.

Compute the second derivative from Equation (47):

$$U_{FF}(F; g, \tau) = \theta\left((2B - \kappa) - 6CF\right) - \lambda(g)\left(S + 6DF^2\right). \tag{19}$$

Proposition 1 (Sufficient condition for uniqueness)**.** *If*

$$\theta(2B(\tau) - \kappa) \;\leq\; \lambda(g)S(g, \tau), \tag{20}$$

then $U_{FF}(F; g, \tau) < 0$ for all $F \geq 0$. Hence the optimal inflow $F^\star(g, \tau)$ is unique.

Proof. Step 1: Compute the second derivative: We differentiate the first-order condition $U_F(F)$ with respect to F.

$$U_F(F) = \underbrace{\theta A}_{\text{const}} + \theta(2B - \kappa)F - 3\theta CF^2 - \lambda SF - 2\lambda DF^3$$

Differentiating term by term:

$$\begin{aligned} U_{FF}(F) &= \frac{d}{dF}[\theta(2B-\kappa)F] - \frac{d}{dF}[3\theta CF^2] - \frac{d}{dF}[\lambda SF] - \frac{d}{dF}[2\lambda DF^3] \\ &= \theta(2B-\kappa) - 6\theta CF - \lambda S - 6\lambda DF^2 \end{aligned}$$

Grouping the constant terms (independent of F) and the F-dependent terms:

$$U_{FF}(F) = [\theta(2B-\kappa) - \lambda S] - (6\theta CF + 6\lambda DF^2) \tag{21}$$

Step 2: Sign analysis: We evaluate the sign of $U_{FF}(F)$ under the condition given in the Proposition.

First, we analyze the constant term: the hypothesis of the proposition is:

$$\theta(2B-\kappa) \leq \lambda S$$

Rearranging this inequality, we get:

$$\theta(2B-\kappa) - \lambda S \leq 0$$

Let $K_0 = \theta(2B-\kappa) - \lambda S$. Thus, $K_0 \leq 0$.

Second, analyze the F-dependent terms: consider the term $-(6\theta CF + 6\lambda DF^2)$. From the model definitions: i) $\theta > 0$ (preference parameter), ii) $\lambda > 0$ (risk aversion), iii) $C = (1-\tau)RJ\eta$. Since crash parameters $J, \eta > 0$, we have $C > 0$, iv) $D = (1-\tau)^2R^2J^2\eta$. Since $J, \eta > 0$, we have $D > 0$, v) the investment level is positive, $F > 0$.

Therefore, for any $F > 0$:

$$6\theta CF > 0 \quad \text{and} \quad 6\lambda DF^2 > 0$$

This implies:

$$-(6\theta CF + 6\lambda DF^2) < 0$$

Conclusion: Combining the terms for any $F > 0$:

$$U_{FF}(F) = \underbrace{K_0}_{\leq 0} + \underbrace{-(6\theta CF + 6\lambda DF^2)}_{<0}$$

$$U_{FF}(F) < 0$$

Since the second derivative is strictly negative for all $F > 0$, the objective function $U(F)$ is strictly concave on the domain $(0, \infty)$.

A strictly concave function defined on a convex set can have at most one global maximum. Lemma 1 established that a maximum exists; strictly concavity ensures that this maximum is unique. □

2.5.2 Closed-form solution of the cubic

Write Equation (47) as:

$$a_3F^3 + a_2F^2 + a_1F + a_0 = 0, \tag{22}$$

with coefficients:

$$a_3 = 2\lambda(g)D(\tau), \quad a_2 = 3\theta C(g,\tau), \quad a_1 = \lambda(g)S(g,\tau) - \theta(2B(\tau) - \kappa), \quad a_0 = -\theta A(g,\tau). \tag{23}$$

Note $a_3 > 0$, $a_2 > 0$ and $a_0 < 0$ when $A > 0$.

Define the depressed-cubic transformation:

$$F = y - \frac{a_2}{3a_3}. \tag{24}$$

Substituting Equation (24) into Equation (22) yields:

$$y^3 + py + q = 0, \tag{25}$$

where

$$p = \frac{3a_3a_1 - a_2^2}{3a_3^2}, \qquad q = \frac{2a_2^3 - 9a_3a_2a_1 + 27a_3^2a_0}{27a_3^3}. \tag{26}$$

The discriminant is:

$$\Delta = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3. \tag{27}$$

Proposition 2 (Characterization of the real roots)**.** *The depressed cubic Equation (25) has:*

- *exactly one real root if* $\Delta > 0$*;*

- *three distinct real roots if* $\Delta < 0$*;*
- *multiple real roots if* $\Delta = 0$.

In all cases, any real root y maps to a real root F via Equation (24). *Any optimal inflow must be a nonnegative real root satisfying the second-order condition* $U_{FF}(F) < 0$.

Proof. Step 1: Reduction to depressed cubic: We apply the standard Tschirnhaus transformation to eliminate the quadratic term (F^2) from the general cubic equation. Let:

$$F = y - \frac{a_2}{3a_3} \tag{28}$$

Substituting this into the general cubic equation $a_3F^3 + a_2F^2 + a_1F + a_0 = 0$:

$$a_3\left(y - \frac{a_2}{3a_3}\right)^3 + a_2\left(y - \frac{a_2}{3a_3}\right)^2 + a_1\left(y - \frac{a_2}{3a_3}\right) + a_0 = 0$$

Expanding the terms (algebraic verification): 1) The coefficient of y^3 is a_3. To normalize, we will divide the entire equation by a_3 later. 2) The coefficient of y^2 becomes zero by design: $3a_3(-\frac{a_2}{3a_3}) + a_2 = -a_2 + a_2 = 0$.

After expansion and simplifying, we divide by a_3 to obtain the monic form:

$$y^3 + py + q = 0 \tag{29}$$

where the coefficients p and q are given by the standard reduction formulas:

$$p = \frac{3a_3a_1 - a_2^2}{3a_3^2} \tag{30}$$

$$q = \frac{2a_2^3 - 9a_3a_2a_1 + 27a_3^2a_0}{27a_3^3} \tag{31}$$

Step 2: Discriminant analysis: The nature of the roots of the depressed cubic equation $y^3 + py + q = 0$ is determined by its discriminant, defined as:

$$\Delta = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 \tag{32}$$

Based on the algebraic theory of cubic equations (Cardano's method), the sign of Δ determines the roots as follows:

1. Case $\Delta > 0$: The formula for the roots involves the sum of two cube roots. One cube root

is real, and the other two involve complex numbers that do not cancel out to form real numbers (except for the principal root).

- Result: One real root and two complex conjugate roots.

2. Case $\Delta < 0$: This is the "casus irreducibilis." The roots can be expressed trigonometrically.

   - Result: Three distinct real roots.

3. Case $\Delta = 0$: The roots are real, and there is multiplicity.

   - If $p = q = 0$, there is one triple real root at $y = 0$.
   - If $p \neq 0$, there is one single real root and one double real root.
   - Result: Real roots with multiplicity.

Step 3: Mapping back to investment F: Since the coefficients $a_0, \ldots, a_3$ are real, the translation term $\frac{a_2}{3a_3}$ is real. Therefore: i) a real solution y corresponds uniquely to a real solution $F = y - \frac{a_2}{3a_3}$; ii) complex solutions y correspond to complex solutions F (which are economically meaningless).

Step 4: Optimality: The first-order condition $U_F(F) = 0$ is necessary but not sufficient for a maximum. A valid candidate for the optimal investment F^* must satisfy two additional physical/economic constraints: 1) Feasibility: $F^* \geq 0$; 2) Maximality: the second-order condition must hold, i.e., $U_{FF}(F^*) < 0$.

If $\Delta < 0$, yielding three real roots, the global maximum is found by checking the utility level $U(F)$ at the feasible roots that satisfy the second-order condition, as well as checking the boundary $F = 0$ (though Lemma 1 typically rules out $F = 0$ if $A > 0$). □

When $\Delta > 0$, the unique real root is:

$$y^\star = \sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}}, \qquad F^\star = y^\star - \frac{a_2}{3a_3}. \tag{33}$$

When $\Delta < 0$, a convenient trigonometric representation is:

$$y_k = 2\sqrt{-\frac{p}{3}}\cos\left(\frac{1}{3}\arccos\left(\frac{3q}{2p}\sqrt{-\frac{3}{p}}\right) - \frac{2\pi k}{3}\right), \quad k = 0, 1, 2, \tag{34}$$

and $F_k = y_k - \frac{a_2}{3a_3}$. One then picks the economically relevant $F_k \geq 0$ satisfying the SOC.

2.6 Risk premium, tail risk, and amplification

2.6.1 A "risk-price" analogue

One can summarize incentives by a risk premium price ρ (mean per unit of risk). Here, a natural analogue is:

$$\rho(F, g, \tau) \equiv \frac{\mu_X(F, g, \tau)}{\sigma_X(F, g, \tau)}. \tag{35}$$

Using (16),

$$\rho(F, g, \tau) = \frac{A(g, \tau) + B(\tau)F - C(g, \tau)F^2}{\sqrt{S(g, \tau) + D(\tau)F^2}}. \tag{36}$$

The numerator is the carry net of endogenous crash losses; the denominator is total risk (normal volatility plus jump risk that rises with inflows).

Lemma 2 (Endogenous tail risk lowers the risk price at high inflows)**.** *Fix* (g, τ)*. If* $C(g, \tau) > 0$ *and* $D(\tau) > 0$*, then* $\rho(F, g, \tau) \to -\infty$ *as* $F \to \infty$*. Moreover, there exists* $\bar{F}_\rho > 0$ *such that* $\rho(F, g, \tau)$ *is strictly decreasing for all* $F \geq \bar{F}_\rho$*.*

Proof. Part 1: Asymptotic behavior (limit): We examine the limit of the ratio as $F \to \infty$. We divide both the numerator and the denominator by the highest power of F present in the denominator, which is $\sqrt{F^2} = F$.

$$\rho(F) = \frac{A + BF - CF^2}{\sqrt{S + DF^2}} = \frac{F(\frac{A}{F} + B - CF)}{F\sqrt{\frac{S}{F^2} + D}}$$

Canceling F for $F > 0$:

$$\rho(F) = \frac{\frac{A}{F} + B - CF}{\sqrt{\frac{S}{F^2} + D}}$$

Now, take the limit as $F \to \infty$: i) the numerator behaves as: $0 + B - \infty \to -\infty$ (since $C > 0$); ii) the denominator behaves as: $\sqrt{0 + D} = \sqrt{D}$ (a positive constant, since $D > 0$). Thus:

$$\lim_{F \to \infty} \rho(F) = \frac{-\infty}{\sqrt{D}} = -\infty$$

This proves the first part of the lemma. Economically, the expected return declines quadratically (due to crash probability ηF^2), while volatility increases only linearly (square root of F^2). The crash drag eventually dominates.

Part 2: Monotonicity for large inflows: We calculate the first derivative of $\rho(F)$ with respect to

F using the quotient rule. Let $N(F) = A + BF - CF^2$ and $G(F) = \sqrt{S + DF^2}$.

$$\rho'(F) = \frac{N'(F)G(F) - N(F)G'(F)}{[G(F)]^2}$$

Calculate the components:

$$N'(F) = B - 2CF$$
$$G'(F) = \frac{d}{dF}(S + DF^2)^{1/2} = \frac{1}{2}(S + DF^2)^{-1/2} \cdot 2DF = \frac{DF}{\sqrt{S + DF^2}}$$

Substitute these into the expression for $\rho'(F)$:

$$\rho'(F) = \frac{(B - 2CF)\sqrt{S + DF^2} - (A + BF - CF^2)\frac{DF}{\sqrt{S+DF^2}}}{S + DF^2}$$

To determine the sign, we focus on the numerator of this fraction. We multiply the numerator and denominator by $\sqrt{S + DF^2}$ to clear the fraction in the numerator term. The sign of $\rho'(F)$ is determined entirely by the sign of the resulting numerator, we call it $\mathcal{N}(F)$:

$$\begin{aligned}\mathcal{N}(F) &= (B - 2CF)(S + DF^2) - (A + BF - CF^2)DF \\ &= (BS + BDF^2 - 2CFS - 2CDF^3) - (ADF + BDF^2 - CDF^3)\end{aligned}$$

Group terms by powers of F: i) F^3 terms: $-2CDF^3 - (-CDF^3) = -CDF^3$; ii) F^2 terms: $BDF^2 - BDF^2 = 0$; iii) F^1 terms: $-2CFS - ADF = -(2CS + AD)F$; and iv) constant terms: BS.

Thus, the numerator simplifies to a cubic polynomial:

$$\mathcal{N}(F) = -CDF^3 - (2CS + AD)F + BS$$

Since $C > 0$ and $D > 0$, the coefficient of the leading term (F^3) is strictly negative ($-CD < 0$).

Therefore, as $F \to \infty$, $\mathcal{N}(F) \to -\infty$. This implies that there exists some threshold $\overline{F}_\rho$ such that for all $F > \overline{F}_\rho$, $\mathcal{N}(F) < 0$.

Since the denominator of $\rho'(F)$ is always positive, it follows that $\rho'(F) < 0$ for all $F > \overline{F}_\rho$. □

2.6.2 Tail risk measure

Define the tail event as the crash state $C = 1$. Conditional tail probability is simply:

$$\mathcal{T}(F, g) \equiv \mathbb{P}(C = 1 \mid F, g) = \pi(F, g) = \bar{\pi}(g) + \eta F^2. \tag{37}$$

Proposition 3 (Inflow elasticity of crash (tail) probability)**.** *Define the tail-event probability (FX crash probability) as*

$$\mathcal{T}(F,g) \equiv \mathbb{P}(C=1 \mid F,g) = \pi(F,g) = \bar{\pi}(g) + \eta F^2,$$

and define the inflow elasticity of tail risk as the log–log derivative

$$\varepsilon_{\mathcal{T},F}(F,g) \;\equiv\; \frac{\partial \log \mathcal{T}(F,g)}{\partial \log F}.$$

Then

$$\varepsilon_{\mathcal{T},F}(F,g) = \frac{2\eta F^2}{\bar{\pi}(g) + \eta F^2}. \tag{38}$$

In particular, if inflows are large enough that the endogenous component dominates the baseline term ($\eta F^2 \gg \bar{\pi}(g)$), then

$$\varepsilon_{\mathcal{T},F}(F,g) \to 2.$$

Equivalently, in that region the crash probability scales approximately like $\mathcal{T}(F,g) \approx \eta F^2$, so a 1% increase in F raises tail risk by about 2%.

Proof. Step 1: Compute the marginal change in tail risk: First, we differentiate the tail probability function $\mathcal{T}(F,g)$ with respect to the inflow F. Treating g as a constant:

$$\frac{\partial \mathcal{T}(F,g)}{\partial F} = \frac{\partial}{\partial F}\left(\overline{\pi}(g) + \eta F^2\right)$$

Since $\overline{\pi}(g)$ is constant with respect to F:

$$\frac{\partial \mathcal{T}(F,g)}{\partial F} = 0 + 2\eta F = 2\eta F$$

Step 2: Compute the elasticity: We substitute the derivative and the original function into the definition of elasticity:

$$\epsilon_{\mathcal{T},F}(F,g) = \frac{F}{\mathcal{T}(F,g)} \cdot \frac{\partial \mathcal{T}(F,g)}{\partial F}$$

Substituting $\mathcal{T} = \overline{\pi} + \eta F^2$ and $\frac{\partial \mathcal{T}}{\partial F} = 2\eta F$:

$$\epsilon_{\mathcal{T},F}(F,g) = \frac{F}{\overline{\pi}(g) + \eta F^2} \cdot (2\eta F)$$

Multiplying the numerator terms:

$$\epsilon_{\mathcal{T},F}(F,g) = \frac{2\eta F^2}{\overline{\pi}(g) + \eta F^2} \tag{39}$$

This establishes the main formula of the proposition.

Step 3: Analyze the limit (dominant endogenous risk): We examine the behavior of this elasticity when the endogenous component of risk (ηF^2) becomes very large relative to the baseline risk ($\overline{\pi}(g)$).

Assume $\eta F^2 \gg \overline{\pi}(g)$. We can approximate the denominator:

$$\overline{\pi}(g) + \eta F^2 \approx \eta F^2$$

Substituting this into the elasticity formula:

$$\epsilon_{\mathcal{T},F}(F,g) \approx \frac{2\eta F^2}{\eta F^2} = 2$$

Alternatively, we can take the limit formally as $F \to \infty$ (assuming the model holds for large F):

$$\lim_{F\to\infty} \epsilon_{\mathcal{T},F}(F,g) = \lim_{F\to\infty} \frac{2\eta F^2}{\overline{\pi} + \eta F^2} = \lim_{F\to\infty} \frac{2\eta}{\frac{\overline{\pi}}{F^2} + \eta} = \frac{2\eta}{0 + \eta} = 2$$

Economic interpretation: An elasticity of 2 implies that a 1% increase in capital inflows leads to approximately a 2% increase in the probability of a crash. This "convex" sensitivity arises because the crash probability is a quadratic function of inflows. □

2.6.3 Amplification and multiple equilibria (fixed-point interpretation)

Equation (47) already embeds two opposing feedbacks: i) *carry amplification*: if $\chi > 0$, inflows raise expected appreciation, increasing expected carry (term BF); ii) *fragility amplification*: inflows raise crash probability, lowering expected carry (term $-CF^2$) and increasing risk (term DF^2). When the carry amplification dominates locally but fragility dominates globally, the quartic objective can become non-concave, delivering multiple local extrema and thus multiple candidate inflow "regimes".

A crisp sufficient condition for possible multiplicity is that the objective is locally convex near $F = 0$:

$$U_{FF}(0; g, \tau) = \theta(2B(\tau) - \kappa) - \lambda(g)S(g,\tau) \; > \; 0. \tag{40}$$

This is the opposite of the uniqueness condition in Proposition 1.

Theorem 1 (Sudden-stop geometry via non-concavity)**.** *Suppose $A(g,\tau) > 0$ and Equation (40) holds, while $D(\tau) > 0$ and $\lambda(g) > 0$. Then $U(F; g, \tau)$ is increasing at $F = 0$, locally convex near 0, and eventually tends to $-\infty$. Hence there exist parameter values for which*

$U(F; g, \tau)$ admits at least two critical points on $(0, \infty)$, and therefore the FOC (47) can have multiple positive roots. For such parameter values, small changes in (g, τ) can move the global maximizer discontinuously from a high-inflow root to a low-inflow root.

Proof. We analyze the geometry of the investor's objective function $U(F)$ to determine the possibility of multiple equilibria. Recall the quartic objective function derived in Equation (17):

$$U(F) = \theta \left(AF + BF^2 - CF^3 - \frac{\kappa}{2}F^2 \right) - \frac{\lambda}{2}F^2(S + DF^2)$$

The first-order condition (FOC) is given by $U'(F) = 0$:

$$U'(F) = \theta A + [\theta(2B - \kappa) - \lambda S]\, F - 3\theta CF^2 - 2\lambda DF^3 \tag{41}$$

The second derivative is:

$$U''(F) = [\theta(2B - \kappa) - \lambda S] - 6\theta CF - 6\lambda DF^2 \tag{42}$$

We then analyze the behavior of the function and its slope at the starting point of investment, $F = 0$.

1. Slope ($U'(0)$): Evaluating the first derivative (Equation (41)) at $F = 0$:

$$U'(0) = \theta A + 0 - 0 - 0 = \theta A$$

Since the theorem assumes $A > 0$ (positive expected carry) and $\theta > 0$, we have:

$$U'(0) > 0$$

This implies that the utility function is strictly increasing at the origin. The investor has an incentive to start investing.

2. Curvature ($U''(0)$): Evaluating the second derivative (Equation (42)) at $F = 0$:

$$U''(0) = \theta(2B - \kappa) - \lambda S - 0 - 0 = \theta(2B - \kappa) - \lambda S$$

The theorem explicitly assumes condition (40), which states that this term is positive. Therefore:

$$U''(0) > 0$$

This implies that the utility function is locally convex near zero.

Economic interpretation: In this region, the “carry amplification” (the feedback where inflows appreciate the currency and increase returns, driven by B) dominates the “fragility” and risk aversion effects. The marginal utility of investing is increasing for small positions.

Step 2: Geometry at infinity ($F \to \infty$): We examine the tail behavior. From Lemma 1, we know the term with the highest power of F in $U(F)$ is $-\frac{\lambda D}{2}F^4$. Since $\lambda > 0$ and $D > 0$:

$$\lim_{F\to\infty} U(F) = -\infty$$

Consequently, the derivative must eventually become negative:

$$\lim_{F\to\infty} U'(F) = -\infty$$

Step 3: Intermediate geometry and multiplicity: We combine the findings: i) the function starts at $U(0) = 0$ with a positive slope ($U'(0) > 0$); ii) the slope is initially increasing ($U''(0) > 0$), meaning the function is becoming steeper (convex); iii) eventually, the function must turn around and crash to $-\infty$.

For a smooth function to transition from being increasing and convex (accelerating upwards) to approaching $-\infty$, it must undergo significant curvature changes: i) it must transition from convex to concave (an inflection point); ii) it must reach a local maximum (where $U' = 0$).

Under the specific condition where the convexity at the origin is strong enough, the function develops a non-trivial “S-shape” (convex then concave) before declining. This geometry allows for complex dynamics where the global maximum can shift discontinuously between a high-inflow solution (on the far side of the hill) and a low-inflow solution (or corner solution) as parameters change.

Thus, under condition (40), the standard uniqueness guarantee fails, and the objective function admits the geometry necessary for multiple equilibria (sudden stops). □

Corollary 1 (Macroprudential wedges can remove multiplicity)**.** *Fix g. If τ is increased so that $\theta(2B(\tau) - \kappa) \le \lambda(g)S(g,\tau)$ holds (as in Proposition 1), then the inflow optimum is unique and the discontinuous switching described in Theorem 1 is ruled out.*

Proof. Step 1: Link to global concavity: Assume the policy wedge τ is set such that the condition $\theta(2B(\tau) - \kappa) \le \lambda(g)S(g,\tau)$ is satisfied. From the proof of Proposition 1, we know that this condition implies:

$$U_{FF}(F; g, \tau) < 0 \quad \text{for all } F \ge 0$$

This means the investor’s objective function $U(F)$ becomes globally strictly concave over the entire domain of feasible inflows.

Step 2: Uniqueness of the optimum: A strictly concave function defined on a convex set (the non-negative real line $[0, \infty)$) can have at most one global maximum. Since Lemma 1 ensures that at least one maximum exists (provided $A > 0$), global strict concavity guarantees that this maximum is unique.

Step 3: Ruling out discontinuous switching: Theorem 1 describes a scenario where small changes in (g, τ) can move the global maximizer discontinuously. This behavior relies on the existence of multiple local maxima separated by a local minimum (the "S-shape" geometry arising from non-concavity).

By enforcing the condition in Step 1, the geometry of $U(F)$ is restricted to be a single "hill" (inverted U-shape). Specifically, i) there is only one peak; ii) the location of this single peak, $F^*(g, \tau)$, varies continuously with the parameters (g, τ) (by the Implicit Function Theorem applied to the strictly concave FOC).

Therefore, the possibility of jumping between distinct roots (discontinuous switching) is eliminated. The macroprudential tax τ successfully stabilizes the market by dampening the "carry amplification" mechanism (represented by $B(\tau)$) sufficiently to ensure the risk-aversion forces dominate. □

2.7 Comparative statics and policy predictions

This section records sharp derivatives. The key tool is the implicit function theorem applied to the FOC (47).

Let

$$\Psi(F; g, \tau) \equiv U_F(F; g, \tau).$$

An interior optimum satisfies $\Psi(F^\star; g, \tau) = 0$ and $U_{FF}(F^\star; g, \tau) < 0$.

2.7.1 Effect of the macroprudential wedge

Proposition 4 (Higher τ lowers inflows and tail risk)**.** *Assume $F^\star(g, \tau)$ is an interior optimum with $U_{FF}(F^\star; g, \tau) < 0$. Then*

$$\frac{\partial F^\star}{\partial \tau} < 0, \qquad \frac{\partial \mathcal{T}(F^\star, g)}{\partial \tau} < 0. \tag{43}$$

Proof. Part 1: Effect on inflows (F^*): We apply the Implicit Function Theorem to the condition

$\Psi(F^*; g, \tau) = 0$:

$$\frac{\partial F^*}{\partial \tau} = -\frac{\Psi_\tau(F^*; g, \tau)}{\Psi_F(F^*; g, \tau)} \tag{44}$$

Since the denominator $\Psi_F = U_{FF}$ is strictly negative (by the SOC), the sign of $\frac{\partial F^*}{\partial \tau}$ is the same as the sign of Ψ_τ. We must show that $\Psi_\tau < 0$.

Step 1.1: Structural form of the FOC: Recall the structure of the utility function terms. The expected excess return depends on $(1-\tau)$, while the variance depends on $(1-\tau)^2$. We can rewrite the FOC $\Psi(F)$ compactly by grouping the terms associated with the marginal return and marginal risk.

Let $\mathcal{M}(F)$ be the gross marginal return component (before tax and funding costs) and $\mathcal{V}_{marg}(F)$ be the marginal contribution to variance. The FOC can be expressed as:

$$\Psi(F, \tau) = \theta\left[(1-\tau)\mathcal{M}(F) - R^* - \kappa F\right] - \lambda(1-\tau)^2\mathcal{V}_{marg}(F) = 0 \tag{45}$$

where $\mathcal{M}(F)$ contains terms like $R(1+\mu_e)$, $2R\chi F$, etc; $\mathcal{V}_{marg}(F)$ contains terms like S and DF^3; R^* is the funding cost and κF is the marginal balance sheet cost.

Step 1.2: Differentiating w.r.t. τ: Compute the partial derivative of Ψ with respect to τ:

$$\begin{aligned}
\Psi_\tau &= \frac{\partial}{\partial \tau}\left(\theta(1-\tau)\mathcal{M}(F) - \theta R^* - \theta\kappa F - \lambda(1-\tau)^2\mathcal{V}_{marg}(F)\right) \\
&= \theta(-\mathcal{M}(F)) - \lambda \cdot 2(1-\tau)(-1)\mathcal{V}_{marg}(F) \\
&= -\theta\mathcal{M}(F) + 2\lambda(1-\tau)\mathcal{V}_{marg}(F)
\end{aligned}$$

Step 1.3: Substitution from FOC: From the FOC equation (45), we can solve for the risk term $\lambda(1-\tau)\mathcal{V}_{marg}(F)$:

$$\begin{aligned}
\lambda(1-\tau)^2\mathcal{V}_{marg}(F) &= \theta\left[(1-\tau)\mathcal{M}(F) - R^* - \kappa F\right] \\
\implies \lambda(1-\tau)\mathcal{V}_{marg}(F) &= \frac{\theta}{1-\tau}\left[(1-\tau)\mathcal{M}(F) - R^* - \kappa F\right]
\end{aligned}$$

Substitute this back into the expression for Ψ_τ:

$$\begin{aligned}
\Psi_\tau &= -\theta\mathcal{M}(F) + 2\left(\frac{\theta}{1-\tau}\left[(1-\tau)\mathcal{M}(F) - R^* - \kappa F\right]\right) \\
&= \frac{\theta}{1-\tau}\left[-(1-\tau)\mathcal{M}(F) + 2(1-\tau)\mathcal{M}(F) - 2R^* - 2\kappa F\right] \\
&= \frac{\theta}{1-\tau}\left[(1-\tau)\mathcal{M}(F) - 2R^* - 2\kappa F\right]
\end{aligned}$$

Step 1.4: Determining the sign: We can substitute $(1-\tau)\mathcal{M}(F)$ using the FOC again. Since

$\Psi(F) = 0$:

$$(1-\tau)\mathcal{M}(F) = R^* + \kappa F + \underbrace{\frac{\lambda}{\theta}(1-\tau)^2 \mathcal{V}_{marg}(F)}_{\text{Risk Premium}>0}$$

Substitute this into the bracket for Ψ_τ:

$$\begin{aligned}\Psi_\tau &= \frac{\theta}{1-\tau}\left[(R^* + \kappa F + \text{Risk Premium}) - 2R^* - 2\kappa F\right] \\ &= \frac{\theta}{1-\tau}\left[\text{Risk Premium} - R^* - \kappa F\right]\end{aligned}$$

In standard financial models, the equilibrium risk premium (which is of the order of a few percentage points) is significantly smaller than the gross funding cost R^* (which is $1 + r^* > 1$). Therefore:

$$\text{Risk Premium} - R^* < 0$$

Since $\kappa F \geq 0$, the entire term in the bracket is strictly negative.

$$\Psi_\tau < 0$$

Thus, from the Implicit Function Theorem:

$$\frac{\partial F^*}{\partial \tau} = -\frac{\Psi_\tau}{\Psi_F} = -\frac{(-)}{(-)} < 0$$

Part 2: Effect on tail risk ($\mathcal{T}$): Recall from Proposition 3 that the crash probability is given by:

$$\mathcal{T}(F, g) = \overline{\pi}(g) + \eta F^2$$

Differentiating with respect to τ using the chain rule:

$$\frac{\partial \mathcal{T}}{\partial \tau} = \frac{\partial \mathcal{T}}{\partial F} \cdot \frac{\partial F^*}{\partial \tau}$$

We know: 1) $\frac{\partial \mathcal{T}}{\partial F} = 2\eta F > 0$ (tail risk increases with inflows); 2) $\frac{\partial F^*}{\partial \tau} < 0$ (from Part 1).

Therefore:

$$\frac{\partial \mathcal{T}}{\partial \tau} < 0$$

A higher macroprudential wedge reduces the equilibrium inflow, which directly reduces the endogenous probability of a crash. □

2.8 Global cycle comparative statics

To connect to global financial cycle, impose standard sign restrictions:

Assumption 3 (Global cycle monotonicities)**.** *For higher g (easier global conditions):*

$$\frac{dR^*(g)}{dg} < 0, \quad \frac{d\lambda(g)}{dg} < 0, \quad \frac{d\bar{\pi}(g)}{dg} < 0, \quad \frac{d\sigma_\varepsilon(g)}{dg} \leq 0, \quad \frac{d\mu_e(g)}{dg} \geq 0.$$

Proposition 5 (Easier global conditions raise inflows)**.** *Under Assumption 3, and at any interior optimum with $U_{FF}(F^\star; g, \tau) < 0$, we have:*

$$\frac{\partial F^\star}{\partial g} > 0 \quad \textit{for parameter regions where baseline carry } A(g,\tau) \textit{ rises in } g. \tag{46}$$

Proof. We define $\Psi(F; g, \tau)$ as the First-Order Condition (FOC) for the investor's optimization problem. From Equation (47):

$$\Psi(F, g) = \theta\left(A(g) + (2B - \kappa)F - 3C(g)F^2\right) - \lambda(g)\left(S(g)F + 2DF^3\right) = 0 \tag{47}$$

(We suppress τ as it is held fixed for this proposition).

An interior optimum F^* satisfies $\Psi(F^*, g) = 0$ and the Second-Order Condition (SOC) $\Psi_F < 0$.

We invoke Assumption 3 (Global cycle monotonicities), which states that for higher g (easier conditions): i) funding cost falls: $R^*_g < 0$; ii) risk aversion falls: $\lambda_g < 0$; iii) baseline crash probability falls: $\bar{\pi}_g < 0$; iv) baseline volatility falls or is constant: $(\sigma^2_\epsilon)_g \leq 0$; v) baseline appreciation rises or is constant: $\mu_{e,g} \geq 0$.

We apply the Implicit Function Theorem to the condition $\Psi(F^*, g) = 0$:

$$\frac{\partial F^*}{\partial g} = -\frac{\Psi_g(F^*, g)}{\Psi_F(F^*, g)} \tag{48}$$

Since the denominator Ψ_F is strictly negative (by the SOC), the sign of $\frac{\partial F^*}{\partial g}$ is the same as the sign of Ψ_g. We must show that $\Psi_g > 0$.

Step 1: Expand Ψ_g: Differentiate the FOC expression in Equation (47) with respect to g:

$$\Psi_g = \theta\left(\frac{\partial A}{\partial g} - 3F^2\frac{\partial C}{\partial g}\right) - \left[\frac{\partial \lambda}{\partial g}(SF + 2DF^3) + \lambda(g)\left(\frac{\partial S}{\partial g}F + 2F^3\frac{\partial D}{\partial g}\right)\right]$$

We now analyze the derivatives of the coefficients A, C, S, D with respect to g, based on their definitions (Equations (11)-(15)): 1) Coefficient D: $D = (1-\tau)^2 R^2 J^2 \eta$. Since η and J are constant parameters (fragility is structural), D does not depend on g.

$$\frac{\partial D}{\partial g} = 0$$

2. Coefficient C: $C = (1-\tau)RJ\eta$. Similarly, C is constant with respect to g.

$$\frac{\partial C}{\partial g} = 0$$

3. Coefficient A: $A(g) = (1-\tau)R(1+\mu_e(g) - J\overline{\pi}(g)) - R^*(g)$.

$$\frac{\partial A}{\partial g} = (1-\tau)R\left(\frac{\partial \mu_e}{\partial g} - J\frac{\partial \overline{\pi}}{\partial g}\right) - \frac{\partial R^*}{\partial g}$$

Using Assumption 3: i) $\frac{\partial \mu_e}{\partial g} \geq 0$ (appreciation rises); ii) $-J\frac{\partial \overline{\pi}}{\partial g} > 0$ (crash risk falls, so expected return rises); iii) $-\frac{\partial R^*}{\partial g} > 0$ (funding cost falls). 4. Coefficient S: $S(g) = (1-\tau)^2R^2(\sigma_\epsilon^2(g) + J^2\overline{\pi}(g))$.

$$\frac{\partial S}{\partial g} = (1-\tau)^2R^2\left(\frac{\partial \sigma_\epsilon^2}{\partial g} + J^2\frac{\partial \overline{\pi}}{\partial g}\right)$$

Using Assumption 3: i) $\frac{\partial \sigma_\epsilon^2}{\partial g} \leq 0$ (volatility falls); ii) $\frac{\partial \overline{\pi}}{\partial g} < 0$ (crash probability falls). Thus, $\frac{\partial S}{\partial g} < 0$.

Step 2: Evaluate the sign of Ψ_g: Substitute these results back into the expression for Ψ_g:

$$\Psi_g = \underbrace{\theta\frac{\partial A}{\partial g}}_{>0} - \underbrace{3\theta F^2(0)}_{0} - \underbrace{\frac{\partial \lambda}{\partial g}}_{<0}\underbrace{(SF + 2DF^3)}_{>0} - \lambda(g)\underbrace{\left(\frac{\partial S}{\partial g}F + 0\right)}_{<0}$$

implifying the signs: i) Term 1: $\theta A_g > 0$ (Expected return channel); ii) Term 2: $-\lambda_g(Risk) > 0$ (Since λ_g is negative, $-\lambda_g$ is positive. (lower risk aversion boosts demand); iii) Term 3: $-\lambda S_g F > 0$ (Since S_g is negative, $-\lambda S_g$ is positive. Lower baseline variance boosts demand).

Since all non-zero terms are positive:

$$\Psi_g > 0$$

Conclusion: Applying the Implicit Function Theorem:

$$\frac{\partial F^*}{\partial g} = -\frac{\Psi_g}{\Psi_F} = -\frac{(+)}{(-)} > 0$$

Thus, easier global financial conditions unambiguously increase equilibrium capital inflows. □

3 Conclusion

This paper studies a minimal carry-trade environment in which the probability of an FX crash is endogenous to the scale of portfolio inflows. The key reduced-form assumption is that the crash probability increases convexly with the inflow position, $\pi(F, g) = \bar{\pi}(g) + \eta F^2$, so tail risk is not only a state variable but also an outcome of equilibrium positioning. This single nonlinearity changes the structure of the standard mean–variance portfolio problem in a sharp way: under the rare-crash approximation, the investor's objective becomes quartic in F, and the first-order condition is cubic. The model therefore delivers closed-form characterizations of candidate inflow levels, transparent uniqueness conditions, and a simple mapping from global financial conditions into both the level of inflows and the implied tail probability.

Two results are worth emphasizing. First, because crash risk rises convexly in F, the model distinguishes between periods in which inflows mainly move average returns (through carry and expected appreciation) and periods in which incremental inflows mainly move tail outcomes. In the latter region, the inflow elasticity of tail probability is high: when the endogenous component dominates the baseline term, tail probability scales approximately like F^2 and the log–log elasticity approaches two. Put differently, once the system is far enough into the high-inflow region, small additional inflows can disproportionately raise crash risk even if average returns continue to look attractive.

Second, the same feedback can generate regime-type behavior. When the "carry amplification" channel is strong for small positions but fragility dominates for large positions, the quartic objective can be non-concave and the global maximizer can switch discontinuously between a high-inflow and a low-inflow solution. In this sense, a "sudden stop" can appear as a change in which stationary point is globally preferred, rather than as an imposed occasionally binding constraint. The analytical advantage of this setup is that the conditions for uniqueness versus multiplicity can be stated directly in terms of a few composite coefficients, and the knife-edge between continuous and discontinuous adjustments is explicit.

The model suggests that policy is most valuable when it prevents the economy from drifting into the region where tail risk becomes steep in inflows. This leads to three practical recommendations. First, use simple, position-based tools during inflow booms. A macroprudential wedge on foreigners' after-cost return, τ, reduces equilibrium inflows and therefore reduces the endogenous crash probability. The mechanism is direct: the wedge lowers the private marginal gain from building positions while leaving the fragility channel intact, so the chosen $F^{\star}$ falls. In environments where high-frequency gross inflows are driven by global funding conditions, such a wedge is a blunt but robust way to limit the buildup of crash exposure without needing to precisely identify every micro source of fragility.

Second, make the stance countercyclical in global conditions and local positioning. Because easier global conditions raise the desired inflow position, a natural implementation is a rule-like countercyclical design: tighten the wedge (or margins) when global funding is cheap and risk appetite is high, and relax it when global conditions tighten and private inflows retreat. In practice, this points to conditioning the policy stance on a small set of observables that proxy for the model's state variables and positioning: global risk and funding indicators, nonresident local-currency bond holdings, rapid growth in currency-hedged carry positions, and measures of short-term external funding exposure. The goal is not to fine-tune average inflows, but to avoid entering the region where marginal inflows mainly buy tail risk.

Third, stabilize the regime structure, not only the mean level of inflows. When non-concavity is present, the policy problem is not simply "too much" inflow; it is that the system may sit near a boundary where small shocks change which stationary point is globally optimal. In that case, a policy that restores global concavity is valuable because it removes a source of discontinuous switching. In the model, raising τ dampens the carry amplification channel and can restore uniqueness. An operational reading is that macroprudential policy should be evaluated partly by whether it reduces the likelihood of regime-like jumps in positioning, not just by whether it lowers average inflows in normal times.

Finally, the margin-style instrument in the appendix points to a complementary approach. A VaR-type constraint caps positions precisely when risk rises quickly with F, which is exactly the high-inflow region where endogenous jump risk steepens the variance. This makes margin tools naturally state dependent even when they are set by a simple rule (for example, a fixed confidence level and a time-varying margin tightness).

The framework is intentionally partial equilibrium. Two extensions are immediate. First, one can endogenize local-currency returns (or bond prices) through market clearing, adding a price channel that may strengthen the amplification mechanism. Second, one can introduce domestic borrowers and a welfare criterion to evaluate the trade-off between the benefits of openness and the externality created by position-dependent crash risk. Both extensions would allow a full welfare-based design of macroprudential rules while preserving the main message of the paper: when tail risk is increasing in positions, managing the cycle in risk-taking and leverage is at least as important as managing the cycle in average flows.

References

Adrian, T., and H. S. Shin (2010). Liquidity and leverage. *Journal of Financial Intermediation*, 19(3), 418–437.

Ahmed, S., and A. Zlate (2014). Capital flows to emerging market economies: A brave new world? *Journal of International Money and Finance*, 48(PB), 221–248.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics*, 121(3), 823–866.

Bianchi, J. (2011). Overborrowing and systemic externalities in the business cycle. *American Economic Review*, 101(7), 3400–3426.

Bianchi, J., and E. G. Mendoza (2018). Optimal time-consistent macroprudential policy. *Journal of Political Economy*, 126(2), 588–634.

Broner, F., T. Didier, A. Erce, and S. L. Schmukler (2013). Gross capital flows: Dynamics and crises. *Journal of Monetary Economics*, 60(1), 113–133.

Bruno, V., and H. S. Shin (2015). Capital flows and the risk-taking channel of monetary policy. *Journal of Monetary Economics*, 71(C), 119–132.

Brunnermeier, M. K., S. Nagel, and L. H. Pedersen (2009). Carry trades and currency crashes. In *NBER Macroeconomics Annual 2008*, Vol. 23, 313–347.

Brunnermeier, M. K., and L. H. Pedersen (2009). Market liquidity and funding liquidity. *Review of Financial Studies*, 22(6), 2201–2238.

Burnside, C., M. Eichenbaum, I. Kleshchelski, and S. Rebelo (2011). Do peso problems explain the returns to the carry trade? *Review of Financial Studies*, 24(3), 853–891.

Calvo, G. A. (1998). Capital flows and capital-market crises: The simple economics of sudden stops. *Journal of Applied Economics*, 1, 35–54.

Farhi, E., S. P. Fraiberger, X. Gabaix, R. Rancière, and A. Verdelhan (2009). Crash risk in currency markets. NBER Working Paper, No. 15062.

Farhi, E., and I. Werning (2014). Dilemma not trilemma? Capital controls and exchange rates with volatile capital flows. *IMF Economic Review,* 62(4), 569–605.

Farhi, E., and I. Werning (2016). A theory of macroprudential policies in the presence of nominal rigidities. *Econometrica,* 84(5), 1645–1704.

Forbes, K. J., and F. E. Warnock (2012). Capital flow waves: Surges, stops, flight, and retrenchment. *Journal of International Economics*, 88(2), 235–251.

Fratzscher, M. (2012). Capital flows, push versus pull factors and the global financial crisis. *Journal of International Economics*, 88(2), 341–356.

Geanakoplos, J. (2010). The leverage cycle. In *NBER Macroeconomics Annual 2009*, Vol. 24, 1–65.

Gourinchas, P.-O., and M. Obstfeld (2012). Stories of the twentieth century for the twenty-first. *American Economic Journal: Macroeconomics*, 4(1), 226–265.

Gourio, F. (2012). Disaster risk and business cycles. *American Economic Review*, 102(6), 2734–2766.

Jeanne, O., and A. Korinek (2010). Excessive volatility in capital flows: A Pigouvian taxation approach. *American Economic Review*, 100(2), 403–407.

Jurek, J. W. (2014). Crash-neutral currency carry trades. *Journal of Financial Economics*, 113(3), 325–347.

Korinek, A. (2011). The new economics of prudential capital controls: A research agenda. *IMF Economic Review*, 59, 523–561.

Lustig, H., N. Roussanov, and A. Verdelhan (2011). Common risk factors in currency markets. *Review of Financial Studies*, 24(11), 3731–3777.

n Magud, N. E., C. M. Reinhart, and K. S. Rogoff (2018). Capital controls: Myth and reality—A portfolio balance approach. *Annals of Economics and Finance*, 19(1), 1–47.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91.

Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2012). Carry trades and global foreign exchange volatility. *Journal of Finance*, 67(2), 681–718.

Miranda-Agrippino, S., and H. Rey (2020). U.S. monetary policy and the global financial cycle. *Review of Economic Studies*, 87(6), 2754–2776.

Ostry, J. D., A. R. Ghosh, K. Habermeier, M. Chamon, M. S. Qureshi, and D. B. S. Reinhardt (2010). Capital inflows: The role of controls. IMF Staff Position Note, SPN/10/04.

Ostry, J. D., A. R. Ghosh, M. Chamon, and M. S. Qureshi (2012). Tools for managing financial-stability risks from capital inflows. *Journal of International Economics*, 88(2), 407–421.

Rey, H. (2015). Dilemma not trilemma: The global financial cycle and monetary policy independence. NBER Working Paper, No. 21162.

A A VaR-style Margin Constraint (An Alternative Policy Instrument)

Some macroprudential toolkits operate through margins rather than direct taxes. A simple way to capture this is a Value-at-Risk constraint on the position:

$$\mathrm{VaR}_\alpha(FX) \;\le\; m\,W, \tag{A.1}$$

where $\alpha \in (0, 1)$ is the tail confidence level; $m \in (0, 1)$ is an exogenous margin tightness set by the regulator (lower m is tighter); $W > 0$ is investor wealth (or intermediation capacity).

Under the (common) normal-approximation $\mathrm{VaR}_\alpha(FX) \approx z_\alpha\, F\, \sigma_X(F, g, \tau)$ with $z_\alpha > 0$ the Gaussian quantile, the constraint becomes:

$$F\,\sigma_X(F, g, \tau) \;\le\; \frac{mW}{z_\alpha}. \tag{A.2}$$

Using the approximation $\sigma_X(F, g, \tau) = \sqrt{S + DF^2}$ from (16), Equation (A.2) implies an explicit upper bound:

$$F^2\left(S + DF^2\right) \le \left(\frac{mW}{z_\alpha}\right)^2 \quad\Longrightarrow\quad F \le \bar{F}_{\mathrm{VaR}}(g, \tau, m) \equiv \sqrt{\frac{-S + \sqrt{S^2 + 4D(mW/z_\alpha)^2}}{2D}}. \tag{A.3}$$

The constrained optimum is then

$$F^{\star,\mathrm{VaR}}(g, \tau, m) = \min\left\{F^\star(g, \tau),\ \bar{F}_{\mathrm{VaR}}(g, \tau, m)\right\}.$$

Since $\partial \bar{F}_{\mathrm{VaR}}/\partial m > 0$, tighter margins (lower m) weakly reduce inflows even when the unconstrained solution would be large. This instrument is particularly effective in the high-inflow regime where jump risk makes σ_X steep in F.