A Theoretical Approach to Stablecoin Design via Price Windows: Stability and Reserve Depletion

Katherine Molinet[1[0009-0007-6138-6831]] and Aris Filos-Ratsikas[1[0000-0001-7868-8114]]

University of Edinburgh

Abstract. In this paper, we explore the short- and long-term stability of backed stablecoins offering constant mint and redeem prices to all agents. We refer to such designs as *price window-based,* since the mint and redeem prices constrain the stablecoin's market equilibrium. We show that, without secondary stabilization mechanisms, price window designs cannot achieve both short- and long-term stability unless they are backed by already-stable reserves. In particular, the mechanism faces a tradeoff: either risk eventual reserve depletion through persistent arbitrage by a speculator, or widen the distance between mint and redeem prices enough to disincentivize arbitrage. In the latter case, however, the market price of the stablecoin inherits the volatility of its backing asset, with fluctuations that can be proportional to the backing asset's own volatility.

Keywords: stablecoins · price windows · theoretical framework.

1 Introduction

A stablecoin is a cryptocurrency whose price is "pegged" to some reference asset (typically the U.S. Dollar). Since their emergence in the mid-2010s, a wide range of designs and stabilization mechanisms have been developed. But despite strong incentives and many design attempts, no stablecoin has yet been shown to be both provably stable and fully decentralized. Designs backed by USD, for instance, have been shown to be stable when reserves are well-managed; but the management of off-chain assets requires trust – the very thing cryptocurrencies were created to circumvent. By contrast, cryptocurrency-backed designs – which can be managed by smart contracts instead of human custodians – can avoid this issue, but have struggled with market price volatility. Algorithmic stablecoins seek to avoid a dependence on any external currency, instead controlling market price through regulation of supply and demand, much like a central bank. However, these designs have experienced some of the most catastrophic stablecoin failures to date (e.g., [23]).

Two foundational questions in the field are therefore still open: Can a decentralized, stable coin even exist? And what is the full space of possible stablecoin designs? While our work in this paper represents only a very small part of the complete stablecoin picture, we believe that our approach – in which we use simple mathematical models and minimal assumptions to distinguish between different families of stablecoin designs by the emergent properties of those systems – could be built upon in future work. Our definitions and approaches were crafted with this in mind.

1.1 Our contributions

The scope of this paper is limited to what we call the "price window model," in which the stablecoin mechanism mints and redeems stablecoins at fixed dollar prices, thereby constraining the market price equilibrium to lie between these bounds. This captures familiar USD-backed designs like Tether and USDC, as well as some crypto-backed designs like USDN and Cardano's Djed. Interestingly, these designs are by far the most common among stablecoins backed by USD, but seem to be less preferred among crypto-backed designs. Our work offers a possible explanation as to why.

In Section 2, we suggest that what mathematically distinguishes these designs from other backed stablecoins is the information the mechanism can use to determine the mint and redeem prices it offers. While other designs (such as DAI) can use information about the particular transaction (e.g., the identity and history of the agent initiating the request), price window mechanisms treat all agents and stablecoins *interchangeably.* This property leaves the mechanism vulnerable to a speculator who exploits fluctuations in the price of the backing asset (which we refer to throughout the paper as the "backing coin") to accumulate reserves.[1] Informally, our main result in Section 3 shows that if the stablecoin mechanism is not allowed to use information about the specific transaction, then it cannot effectively guard against such attacks.

Informal Theorem 1. *In the absence of secondary stabilization mechanisms, the long-term sustainability of a price window-based mechanism is only guaranteed when the size of the price window (that is, the difference between the mechanism's mint and redeem prices and, therefore, the rational price interval in the market) scales with the asymptotic volatility of the underlying asset.*

This suggests a fundamental tradeoff between short-term and long-term stability in price window designs over unstable assets. Moreover, our result shows that the higher the volatility of the asset used to back the stablecoin, the higher the potential volatility of the stablecoin. This implies that – at least, for designs which adopt a price window as the primary stabilization mechanism – price stability seems to be largely an inherited property.

We expect Sections 2 and 3 to be most relevant to theorists and researchers looking to define relevant stablecoin design taxonomies and understand emergent properties of different stablecoin designs; but it also has some practical relevance to system designers. For instance, our result provides a theoretical value for

[1] By contrast, mechanisms like DAI with *personalized* redeem prices can guard against this avenue of failure by blocking agents from redeeming stablecoins for more backing coins than they paid to mint – but at the loss of the short-term stability afforded by price windows.

the size of transaction fees needed to prevent long-term depletion of a system's reserves. While this value might be overlarge for systems based on a highly volatile backing coin, it might be more reasonable to implement for mechanisms whose backing coin is close to stable.

Section 4 of this work is designed to answer some of the more practical questions raised by the first part of this work. For instance, in Section 3 we outline a theoretical mechanism by which some abstract speculator could deplete a system's reserves of backing coins, but it's unclear how seriously this risk should be taken. Our claim doesn't address any real-world concerns about whether a real speculator could successfully launch such an attack, and how powerful and well-resourced the speculator would need to be to deplete a system's reserves in a reasonable amount of time. To this end, we move from a theoretical black-box speculator to an explicit model of a speculator and compute a formula for the expected time it would take for the speculator to deplete a system's reserves, given that the backing coin prices are drawn independently at random from some known distribution. We explain how this formula relates the expected depletion time to the volatility of the backing coin, and how the same pattern we noticed in the first section – that the more volatile the backing coin, the less sustainable the system – becomes even more explicit here. We then relax the assumption of independent price draws and simulate the behavior of a speculator over historical Ether and Bitcoin price data, showing that a sufficiently wealthy speculator would have been able to drain the reserves of a stablecoin system backed by each of these currencies within three years.

1.2 Related work

Understanding the source of stability in stablecoin design has been the focus of a growing literature since 2019, e.g., see [25,22,24,15,9,17]. The field of stablecoin literature is inherently interdisciplinary, with sources ranging from empirical and financial analyses [12,16] to theory [6,21] to regulatory and legal reviews [9,3]. While each perspective offers an important layer to the stablecoin narrative, our approach in this paper is theoretical, and so our contribution is best compared against other theoretical papers. In general, the heterogeneity of stablecoins makes it difficult to propose theoretical claims which apply broadly across the design space, and many works respond to this challenge by either focusing on a single type of stablecoin design, or by proposing broader frameworks and taxonomies by which to better understand the options.

For these broader/taxonomic papers, on one end of the spectrum are papers like [5,26,7,8], which offer meaningful qualitative and intuitive observations but can lack mathematical formalism. The work of Bullmann et al. [5] is a paper of this type which has been of particular inspiration to our work, especially the authors' distinction between primary and secondary stabilization mechanisms. Papers which take a more mathematical approach include [17], which classifies backed stablecoins according to the endo/exogeneity of the backing coins; [29], a fascinating paper which uses automata theory to develop a modeling and verification framework which generalizes several types of algorithmic stablecoins;

and a 2024 paper by Potter et al. [27] which, of the taxonomic papers, seems closest to our own. In [27], the authors generalize many known stablecoin designs by their redemption functions and then classify families of designs according to the economic conditions under which they can maintain the $1 peg. However, unlike us, the authors assume that all value is measured in dollars, rather than backing coins. The only failure path for price-window based designs, according to this paper, is through undercollateralization if the price of the backing coin drops sufficiently – something which can only occur in "bad" economic states. Meanwhile, our paper demonstrates how undercollateralization and reserve depletion can occur even in good economic states through a speculator who values backing coins more than dollars. This is a reasonable assumption to make for strong backing coins with high average returns, like Bitcoin or Ether.

With respect to limited-scope theoretical papers, our work differs in its focus on price window-based models over potentially volatile reserves. Since these designs are less commonly seen in practice, it is perhaps understandable that fewer papers have studied them as opposed to, for instance, fiat-backed stablecoins like Tether [24] or overcollateralized designs like DAI [19,18,4,20]. Some notable papers include [19] and [18], which explore deleveraging spirals in stablecoins with DAI-like designs under the influence of an opportunistic speculator. [4] explores undercollateralization risks in BitShares, another overcollateralized, crypto-backed stablecoin. Although this paper explores a different setting from our work, the author's focus on understanding how the mechanism should choose its strategy to prevent some undesired behavior is similar to our own approach. [11] is another notable theoretical paper, but its setup explores systems in which redemption is not guaranteed – instead, investors must sell their stablecoins through the competitive market. The platform defends the peg instead by contracting and expanding the supply of stablecoins.

Several papers, like our own, note a link between the volatility of the backing coin and that of the stablecoin. For instance, [20] shows how unstable backing assets can distort a DAI-like system's mechanism for arbitrage – in particular, the stablecoin tends to trade above (below) its peg when ETH is in a bad (good) state. [27] notes the same behavior as well, though their motivation is wholly different from our own. While we argue that this "inherited (in)stability" occurs through arbitrage incentives, [27] argues that it instead arises from backing coin price discrepancies between agents' expectations and prices offered by the mechanism (due, e.g., to price oracle delays). The more volatile the backing coin, they argue, the more likely it is that these two values will be out-of-sync. In contrast to these papers, in our work we not only show the relationship between the volatility of the stablecoin and backing coin, but we also compute the size of the transaction fees needed to prevent eventual failure.

2 The model

Let $T = \{1, 2, \ldots\} \cong \mathbb{N}$ be a discrete and infinite sequence of *timesteps*, which we will refer to as *time*. In our setup, there are two types of coins: the *tokens* and the *backing coins*. Looking ahead, the former model the *stablecoins*, and the latter model the (possibly volatile) assets that back the stablecoin system, e.g., Ether, BTC, or USD. The value (or *price*) of both tokens and backing coins will be measured in terms of some *fiat currency*; without loss of generality, we assume that this will be the U.S. dollar ($). We will use p_t to denote the dollar price of the backing coin at timestep $t \in T$ (provided, e.g., by an oracle), and assume that $p_t \geq 0$ for all $t \in T$ and that p_t is exogenous to the system.

We define a *token-generating mechanism* $\mathcal{M}$ with which some set of *agents* interact by exchanging backing coins and tokens. The rules of this interaction are governed by the functions α and β, which capture, respectively, how the mechanism *mints* and *redeems* tokens, and which we describe below. Agents are also permitted to trade tokens and backing coins with *each other* in a market. We let d_t denote the equilibrium dollar market price of a token at time t.

In this work (and in the theory of stablecoins more generally), we are interested in token-generating mechanisms which choose interaction functions α and β with the aim of creating the conditions for a *stable* token with market price $d_t = 1$.[2] We call such mechanisms "stablecoin systems" and the tokens they generate "stablecoins," even if they are not always stable in practice. To clarify what we mean by stability, we introduce in this paper a notion which we refer to as "weak ε-stability":

Definition 1. *Let p_t be the price of the backing coin at time t. Given some $\varepsilon \geq 0$, we say that a stablecoin mechanism $\mathcal{M}$ is **ε-stable** over a sequence $(\frac{1}{p_t})_{t\in T}$ if $d_t \in [1-\varepsilon,\, 1+\varepsilon]$ for all $t \in T$.*

Definition 2. *A mechanism $\mathcal{M}$ is **weakly ε-stable** for $\varepsilon \geq 0$ if there exists some non-convergent sequence $(\frac{1}{p_t})_{t\in T}$ over which $\mathcal{M}$ is ε-stable.*

We can conceptualize $(\frac{1}{p_t})_{t\in T}$ as the sequence of prices (in units of backing coin) over time for a stablecoin sold at $1. This notion of "weak stability" has two nice properties. For one, it seems clear that any mechanism which does *not* satisfy this definition must be a fairly ineffective design. For if a design can only achieve stability over a sequence of eventually-convergent (in other words, eventually stable) prices, then the design does not offer any particular long-term improvement; that is, it does not expand the set of price sequences over which stability is achievable. The second nice property of the definition is that it makes a claim about the long-term *sustainability* of a mechanism. It is not sufficient for the mechanism to provide some short-term protection against volatility; to be considered "weakly stable," the mechanism must be able to offer this protection indefinitely over at least one sequence of non-convergent prices. (Note that a design can fail in catastrophic economic conditions (e.g., $p_t = 0$) and still be weakly ε-stable. This is important because otherwise, no backed stablecoin would ever be considered stable.)

[2] This could be any positive constant; but according to convention and without loss of generality, we set it to 1.

2.1 Design taxonomy and price window

Now that we've introduced the primary goal of any stablecoin mechanism, we turn to understanding the various ways these mechanisms try to achieve it. Stablecoin designs are often categorized by the primary and secondary stabilization mechanisms they use.[3] The first major taxonomic distinction we (and many others) make between families of stablecoins designs is whether the stablecoins are *backed* or *unbacked* (also known as "algorithmic"). In this paper, we choose to only focus on the former. For backed stablecoin mechanisms, the value of a stablecoin is derived from agents' ability to redeem their stablecoins to the mechanism for $\beta(\cdot) \approx \$1$ of some currency or asset of value (e.g., USD, ETH, BTC etc.). The currency of choice is what we refer to in this paper as the "backing coin." The mechanism stores its supply of backing coins in its *reserves* R. We let R_t denote the number of backing coins in R at timestep $t \in T$ (with initial state $R_0 > 0$), noting that the value can change at each timestep in response to agents' actions.

We also assume that agents must pay some amount $\frac{\alpha(\cdot)}{p_t}$ of backing coin to the mechanism to purchase/mint a stablecoin at time t.[4] Mathematically, we characterize backed stablecoins by these two operations, minting and redeeming. The mechanism's chosen functions α and β, along with the current backing coin price p_t, indicate how the mechanism maps a request to mint or redeem, respectively, a single stablecoin to some quantity of backing coins received/given in exchange. The information the mechanism uses to decide α and β for each request (i.e., variables upon which α and β depend) gives the next branch in our taxonomy.

Let θ_t represent the state of the system at time t, and let τ represent the particular transaction request. (The state θ, for instance, might include the price of the backing coin and the amount of reserves held in the system, while τ might represent information about the agent initiating the request, as well as information about the stablecoin in question, such as when, by whom, and at what price it was minted.) These two variables together capture all the information upon which the stablecoin mechanism can determine its mint/redeem price functions α and β.[5] From here, we make the following distinction: Any mechanism in which α and β depend solely on the state of the system θ – i.e.,

$$\alpha(\theta_t, \tau_i) = \alpha(\theta_t, \tau_j) \text{ and } \beta(\theta_t, \tau_i) = \beta(\theta_t, \tau_j) \text{ for all } \tau_i, \tau_j \text{ and for any } \theta_t$$

[3] See [5] for a list of some primary and secondary stabilization mechanisms.

[4] While this assumption is not *strictly* necessary theoretically, the burden of any backed stablecoin design, as we shall see, is to maintain sufficient reserves to preserve redeemability. Hence, a mechanism which does not charge users to mint but pays users to redeem is essentially lost before it has even begun.

[5] We might later wish to consider some random variable X, to allow the possibility that α and/or β are stochastic functions; but for now we assume our stablecoin mechanism is deterministic.

– is said to have agent and stablecoin *interchangeability*, since the same mint and redeem prices at any time t are offered to *all* agents and stablecoins, regardless of their identity or history.

These designs include "tokenized" stablecoins like USDC and USDT, as well as some crypto-backed designs like Cardano's Djed and USDN. (Note: The only difference between these two classes of designs, according to our model, is that the backing coins of the former are stable relative to the fiat currency. That is, for tokenized stablecoins, $\frac{1}{p_t} = 1$ for all $t \in T$.) By contrast, backed designs which use information about the specific agent and transaction to determine the redemption price of a stablecoin (like DAI), are considered *non-interchangeable* under this taxonomy. We refer to the stabilization mechanism of interchangeable backed stablecoin designs as a ***"price window"*** (for reasons we will explain shortly), and say that mechanisms which use this approach are following the "price window model" for stablecoin designs.

To understand how such mechanisms work, suppose at time t the mechanism offers some dollar mint/redeem prices $\beta(\theta_t) \leq 1 \leq \alpha(\theta_t)$ to all agents. Agents' ability to redeem stablecoins to the mechanism in exchange for $\$\beta(\theta_t)$-worth of backing coin creates a *price floor* in the market (assuming no market frictions) – for why would any rational agent sell their stablecoin to the market for less than $\$\beta(\theta_t)$, when they could obtain a higher value directly from the mechanism? Similarly, the ability to mint for $\$\alpha(\theta_t)$ creates a *price ceiling* on the market. Together, the price floor and ceiling create a $\$[\beta(\theta_t), \alpha(\theta_t)]$ *price window* around the equilibrium market price d_t of a stablecoin, so long as the mechanism maintains its capacity to mint and redeem stablecoins at these prices.

2.2 Our scope

In what follows, we limit the scope of our analysis to stablecoins which follow a price window design. For the sake of simplicity, we limit this scope even further to price window designs where $\alpha()$ and $\beta()$ are constant with time, since these are the most common designs in practice and can be explored using simpler notation and proofs. We can think of these designs as charging fixed "transaction fees" ε_α, $\varepsilon_\beta \geq 0$ for every mint or redeem request. However, our results – especially in Section 3 – are generalizable to transaction fees ε_α and ε_β which are allowed to vary with time, although we do not include the proof in this paper.[6] We formally define the stablecoin mechanism $\mathcal{M}$ studied in this paper by the following minting and redeeming functions:

- **Minting:** We define a minting price function $\alpha : \mathbb{N} \to \mathbb{R}_{\geq 0}$ which assigns to each time $t \in T \cong \mathbb{N}$ a *mint price*, with units in USD. We write $\alpha(t) = 1 + \varepsilon_\alpha$ for some constant $\varepsilon_\alpha \geq 0$. i.e., $\mathcal{M}$ will always mint a stablecoin for $(1 + \varepsilon_\alpha)$-dollar's-worth of backing coins. In units of backing coin, this corresponds to a price of $\frac{1+\varepsilon_\alpha}{p_t}$ for each time t.

[6] Essentially, when ε_α and ε_β are allowed to vary with time, Theorem 1 still applies, but to $\limsup \varepsilon_\alpha(t)$ and $\limsup \varepsilon_\beta(t)$ rather than ε_α and ε_β.

- **Redeeming:** Similarly, for redeeming we define $\beta : \mathbb{N} \to \mathbb{R}_{\geq 0}$ as follows: for some constant $\varepsilon_\beta \geq 0$, $\beta(t) = 1-\varepsilon_\beta$ if $R_t > \frac{1-\varepsilon_\beta}{p_t}$ and $\beta(t) = R_t p_t$, otherwise. In other words, $\mathcal{M}$ maintains a $(1-\varepsilon_\beta)$-dollar stablecoin sale price whenever it has sufficient reserves; otherwise $\mathcal{M}$ returns as many funds as it can, i.e., the entirety of its reserves. We say that the reserves of $\mathcal{M}$ are eventually *depleted* if there exists some $t \in T$ for which $R_t = 0$.

In this way, and assuming no market frictions, $\mathcal{M}$ maintains a $\$[1-\varepsilon_\beta, 1+\varepsilon_\alpha]$ price window around the market equilibrium price d_t whenever it has sufficient reserves. The first timestep t that $\mathcal{M}$ is unable to maintain the $\$[1-\varepsilon_\beta, 1+\varepsilon_\alpha]$ peg is the time of reserve depletion.

3 General results

Now that we've introduced our model, we explain the failure path to which price-window stablecoin designs are susceptible and show that for small ε, these designs are *not* weakly ε-stable. In fact, reserve depletion will eventually always occur under the sustained effort of a so-called "sensitive speculator" *unless* the sequence of backing coin prices converges, or the size of the transaction fees is proportional to the volatility of the backing coin.

Failure pattern intuition. Consider the lifecycle of a single stablecoin. Minting at backing coin price p_m and redeeming at p_n through a mechanism with a $\$[1-\varepsilon_\beta, 1+\varepsilon_\alpha]$ price window changes the system's reserves by $\frac{1+\varepsilon_\alpha}{p_m} - \frac{1-\varepsilon_\beta}{p_n}$ units of backing coin. If $\frac{1+\varepsilon_\alpha}{p_m} > \frac{1-\varepsilon_\beta}{p_n}$, this impact is positive for the system; if not, it is negative. Over many coins, such effects average out if agents mint and redeem non-strategically. But if a speculator deliberately times minting and redemption to maximize these differences, the system can experience systematic reserve depletion, assuming both that the price of the backing coin is sufficiently volatile relative to the sizes of $\varepsilon_\alpha, \varepsilon_\beta$, and that the speculator is sufficiently skilled at recognizing and acting on backing coin price anomalies. And since any backing coin the speculator gains must be taken from the mechanism's reserves, eventually, sustained arbitrage will exhaust those reserves.

Defining a "sensitive" speculator. Let $\mathcal{M}$ denote a backed stablecoin mechanism with mint and redeem functions $\alpha(t)$ and $\beta(t)$. Define s_t as the total profit (in units of the backing coin) that could be attained by an optimal, omniscient speculator who: 1) begins with a single unit of backing coin, 2) interacts exclusively with $\mathcal{M}$, and 3) knows the entire sequence of backing coin prices $(\frac{1}{p_t})_{t \in T}$ in advance. Then s_t represents the maximum profit achievable up to timestep t over this fixed sequence of prices.

Definition 3. *Suppose a speculator has some number of backing coins n_0 at $t = 0$. We say that a speculator's strategy is $\boldsymbol{\frac{1}{k}}$**-approximately optimal** for some $k \geq 1$ if the total profit r_t obtained by the speculator up to any time t is at least $\frac{1}{k}$th that of the optimal value; i.e., $r_t \geq \frac{n_0}{k} s_t$ for all $t \in T$.*

This definition essentially says that if the mechanism's transaction fees and the sequence of prices offer an opportunity to profit, then the speculator can generally recognize and act on this opportunity, albeit imperfectly. In this way, the definition can be thought of as an assumption about the speculator's "sensitivity" to price fluctuations. We formalize this idea in our next definition.

Definition 4. *A speculator is* ***sensitive*** *if there exists some number $k \geq 1$ such that the speculator's strategy is $\frac{1}{k}$-approximately optimal over any sequence of prices.*

Conveniently, these definitions allow us to construct general results about the asymptotic global dynamics of price window-based stablecoins while making very few assumptions on the internal details of the speculator's strategy or behavior. This leads us to our first main result, the proof of which can be found in the full-text version.

Theorem 1. *A backed stablecoin mechanism $\mathcal{M}$ with finite reserves and constant mint and redeem functions $\alpha(t) = 1 + \varepsilon_\alpha$, $\beta(t) = 1 - \varepsilon_\beta$ for $\varepsilon_\alpha, \varepsilon_\beta \geq 0$ is weakly ε-stable in the presence of any sensitive speculator if and only if ε_α, ε_β satisfy $(1 + \varepsilon_\alpha) \liminf \frac{1}{p_t} - (1 - \varepsilon_\beta) \limsup \frac{1}{p_t} \geq 0$.*

Interpretation Recall that $\limsup \frac{1}{p_t} := \lim_{N\to\infty} \sup \left\{ \frac{1}{p_t} : t > N \right\}$ and that $\liminf \frac{1}{p_t} := \lim_{N\to\infty} \inf \left\{ \frac{1}{p_t} : t > N \right\}$. To interpret Theorem 1, we begin with the simple case of $\varepsilon_\alpha = \varepsilon_\beta = 0$. In this case, our claim implies that reserve depletion occurs if and only if $\liminf \frac{1}{p_t} = \limsup \frac{1}{p_t}$, which, by the squeeze theorem, implies the original sequence $(\frac{1}{p_t})_{t\in T}$ converges. Thus, no price window-based stablecoin mechanism which mints and redeems at \$1 can be weakly 0-stable. The only designs of this type capable of both short- and long-term stability, therefore, are those for which $(\frac{1}{p_t})_{t\in T}$ always converges – for instance, USDT and USDC, where $\frac{1}{p_t} = 1$ for all $t \in T$.

In the more general case, we note that the total profit of sensitive speculators can be scaled to be *arbitrarily small.* As a consequence, the only sequences over which any such speculator is guaranteed to deplete the system's reserves must have infinite total profit. Since each discrete time interval only offers finite profit (even for an optimal speculator), the only way profit can become infinite is in the tail in the price sequence. This explains why our criteria for ε_α and ε_β (and the subsequent proofs) are only concerned with the behavior of the *tail* of the price sequence.

To understand more specifically the role of $\liminf \frac{1}{p_t}$ and $\limsup \frac{1}{p_t}$, we can interpret these values as numbers which capture the range of values in the tail, since lim sup is the tail's least upper bound and lim inf is its greatest lower bound. If $\liminf \frac{1}{p_t}$ and $\limsup \frac{1}{p_t}$ are both finite, then the quantity $\limsup \frac{1}{p_t} - \liminf \frac{1}{p_t}$ represents the range or spread of the tail of the price sequence. This value can be thought of as a measure of the price sequence's asymptotic *volatility*: the larger the difference between $\liminf \frac{1}{p_t}$ and $\limsup \frac{1}{p_t}$, the more volatile the asymptotic behavior of the price sequence.

In this way, charging "transaction fees" can be thought of as a way of flattening the spread of mint and redeem prices (scaling $\limsup \frac{1}{p_t}$ and $\liminf \frac{1}{p_t}$ by $(1-\varepsilon_\beta)$ and $(1+\varepsilon_\alpha)$, respectively) to decrease the range over which a speculator can asymptotically profit. Our claim essentially says that a mechanism can be weakly ε-stable if and only if ε_α and ε_β are large enough to make this "scaled" spread asymptotically tend to zero.

If we set $\varepsilon_\alpha = \varepsilon_\beta = \varepsilon$, then, according to our result, ε must satisfy $\varepsilon \geq \left(\limsup \frac{1}{p_t} - \liminf \frac{1}{p_t}\right) / \left(\limsup \frac{1}{p_t} + \liminf \frac{1}{p_t}\right)$ in order to prevent reserve depletion. Here, we can directly see the relationship between the threshold value of ε and the price sequence's volatility: the two quantities are proportional to one another. (Note that the denominator, $\limsup \frac{1}{p_t} + \liminf \frac{1}{p_t}$, can be interpreted simply as a normalizing factor. It is unrelated to the volatility of the price sequence.) As a result, the range of the price window offered by the mechanism becomes proportional to the spread in the original price sequence $(\frac{1}{p_t})_{t\in T}$. Recall that the key advantage of price windows is their ability to constrain (rational) agents to only trade stablecoins on the market at prices within the offered price window. This is the strategy by which such mechanisms maintain stability. However, our result shows that in order to be sustainable, the size of this price window must essentially be as large as the volatility of the original price sequence. As a result, the market price of stablecoins is allowed to be, proportionately, as volatile as the backing coin against which the system aims to protect.

Of course, transaction fees do help delay a system's time to depletion by adding more reserves to the system as a buffer against arbitrage attacks. However, used in isolation, we conclude that price windows are insufficient to simultaneously ensure both short-term and long-term stability over a volatile backing coin.

4 Case study: A speculator with i.i.d. price draws

The theoretical advantage of the previous section lay in its generality and its ability to capture system-level dynamics with very few assumptions. However, almost necessarily, this generality comes at the cost of overlooking some practical questions that are useful to system designers and managers. In this section, we explore how such specifics can be modeled within the broader framework outlined in Section 3. For instance, how likely is this failure to occur in practice? How powerful must a speculator be, and how long must they persist, to inflict real damage on the system?

To address these questions, we forego our black-box treatment and implement a concrete model of a speculator. Since our earlier claims relied on the existence of a sensitive speculator, here we propose a possible implementation and compute the expected time required for reserve depletion under the assumptions of no transaction fees (i.e., $\varepsilon_\alpha = \varepsilon_\beta = 0$) and independent, identically distributed (i.i.d.) backing coin prices drawn from a known distribution. Note that this section is intended as a high-level summary rather than an in-depth technical analysis. For a full breakdown of the definitions, equations, and results referenced here, we direct the reader to the full-text version, where each concept is laid out in detail.

4.1 Modeling the speculator

In this section, we consider a speculator who has intrinsic value for backing coins but values stablecoins only as a means of obtaining more backing coins. The speculator interacts with a price window-based mechanism $\mathcal{M}$ with $\alpha(t) = \beta(t) = 1$ over a sequence of timesteps, and at each timestep, in response to the revealed backing coin price p_t, either chooses some number Δ_t of stablecoins to buy ($\Delta_t > 0$) or sell ($\Delta_t < 0$); or simply waits until the next timestep ($\Delta_t = 0$). The speculator has some expected utility function u over portfolios of stablecoins and backing coins and chooses Δ_t as the number which maximizes their expected utility.

In our model, we assume the speculator's utility function depends on three parameters: their level of (im)patience δ, their beliefs about backing coin prices, and their risk tolerance Λ. Specifically, $\delta \in [0, 1)$ captures the rate at which the speculator discounts future value, with low δ indicating a patient speculator, and high δ an impatient one. For speculator beliefs about backing coin prices, we assume 1) that backing coin prices are drawn independently at random from some distribution with probability density function $f(x)$, and 2) that $f(x)$ is known to the speculator. Finally, we define the variable $\Lambda \in [0, 1]$ to indicate the speculator's risk tolerance, with $\Lambda = 1$ representing extreme risk aversion and $\Lambda = 0$ representing risk neutrality. To simplify our analytic results, we also assume that the mechanism $\mathcal{M}$ uses a price window design with $\varepsilon = 0$.

Under this model, the speculator's strategy can be seen as a tradeoff: on one hand, they might use all their funds to capitalize on any opportunity for profit, however small, but risk being unable to act when truly exceptional prices occur. On the other hand, they can wait for rare opportunities with exceptional prices, accepting delayed gratification in hopes of obtaining larger profits later.

4.2 Understanding the conditions for (in)stability

In the full-text version, we show how we can solve the speculator's maximization problem at each timestep t and, assuming prices are drawn independently at random from our known distribution, compute a general formula for the expected number n_t of backing coins the speculator owns at any time t for $n_0 > 0$. From here, we can find the expected time to depletion of a mechanism with an initial number R_0 of backing coins in its reserves, though it is not always possible to solve for this time using analytic methods unless $\Lambda = 0$.

Example 1. Suppose backing coin prices are drawn uniformly at random from a normal distribution with $\mu = \sigma^2 = 100$, and suppose the speculator chooses

$\Lambda = 0$ and $\delta = 0.1$. Then setting the ratio $\frac{R_0}{n_0}$ to 100 and computing a_1 according to the formulas in the full-text version gives a reserve depletion which is expected to occur at approximately $k(t) = \log(1+\frac{R_0}{n_0})/\log(a_1) \approx 15.78$, which in turn corresponds to approximately 227 timesteps.[7] We remark that a "timestep" is designated by each time the speculator interacts with the stablecoin mechanism. Assuming, for example, relatively frequent interactions of once every hour, as well as the price updates instructed by our distributional assumptions, this calculation shows that depletion of resources will take place in less than 10 days.

Our formula for the speculator's expected portfolio also allows us to show the analogous "convergence result" to Theorem 1. In the full-text version, we show that the speculator's optimal strategy is to buy stablecoins whenever the backing coin price p_t falls above some threshold price y_2, and to sell whenever p_t falls below some price $y_1 \leq y_2$. From here, we prove that, assuming independent price draws from a distribution with probability density function $f(x)$ and cumulative distribution function $F(x)$, a speculator with $n_0 > 0$ can never deplete a system's reserves if and only if $Y = \frac{\frac{1}{1-F(y_2)}\int_{y_2}^{\infty} x\, f(x)\, dx}{\frac{1}{F(y_1)}\int_{-\infty}^{y_1} x\, f(x)\, dx} = 1$ (or if the speculator is so risk averse that $\Delta_t = 0$ for all $t \in T$). Moreover, the larger the value of Y, the more rapidly the reserves are depleted. We can use a formula from the full-text version to expand the functions $f(x)$ and $F(x)$ in the case of a normal distribution as

$$Y = \frac{\mu + \sigma^2 \frac{f(y_2)}{1-F(y_2)}}{\mu - \sigma^2 \frac{f(y_1)}{F(y_1)}}.$$

From this representation, the relationship between a distribution's volatility and Y is clear: $Y = 1$ if and only if $\sigma^2 = 0$, and the larger the volatility σ^2, the larger the value of Y. This result corresponds nicely to the more general result in Theorem 1, where we showed the relationship between the range of a price sequence's tail and the opportunity for a speculator to deplete the system's reserves. Here, we see that a large value for Y corresponds to a large interval surrounding the tail of the price sequence.

We can see the same relationship in our numerical simulations. Consider, for instance, Figure 1, where we note that the expected time to reserve depletion increases *dramatically* for small values of σ^2 – note that the scale of the y-axis is logarithmic. In the next section, we show that the same behavior arises even when we relax the assumption of independent price draws.

4.3 Simulations

For our results in Section 4.2, we assumed that 1) the price of the backing coin in each timestep is drawn independently at random from some distribution; and 2) the speculator believes that this is the case. While this assumption of

[7] See the full-text version on how to convert from $k(t)$ to the number of expected timesteps.

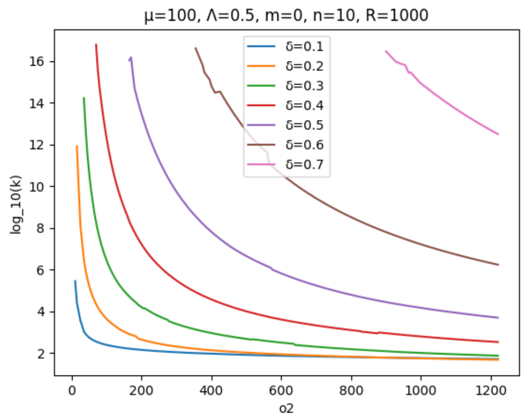

Fig. 1: The log of the expected time to depletion for prices drawn independently at random from a normal distribution with $\mu = 100$ and various values of σ^2, given a speculator with $\delta \in [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7]$, $\Lambda = 0.5$, $n_0 = 10$, and initial reserves $R_0 = 1000$.

independence allows us to produce exact, analytic solutions, in many cases it might not always accurately reflect price fluctuations in reality. In this section, we demonstrate through simulations that the risk of reserve depletion does not depend on the assumption of price independence. Moreover, we find that, as in the case of independent prices, the time it takes for reserves to be depleted decreases as the volatility of the prices grows. Therefore, our simulations convey essentially the same message, just in an environment where the price updates are not independent. This is unsurprising, given the general result shown in Theorem 1.

To achieve this, we coded an adapted version of our speculator and designed a simulator which, given a sequence of backing coin prices[8] and an initial reserve value R_0, repeatedly updates the reserves R_t according to the speculator's response to the revealed backing coin price p_t. The simulation stops either when the reserves have been depleted ($R \leq 0$), or the pre-fixed maximum number of iterations (in our case, 100,000) has been reached. Note that, unlike in the case of independent price draws, the speculator's strategy no longer guarantees that every stablecoin is sold at a higher price than it was bought, nor that reserves are eventually depleted when $\sigma^2 > 0$.

4.4 Simulating with real-world data

To demonstrate that, without any interventions or contingency mechanisms, reserve depletion is nevertheless a very real possibility in real-world stablecoin sys-

[8] This could be a fixed dataset; or could be constructed in real-time e.g., by repeatedly querying some random process. We use both approaches in different simulations.

Initial endowment of $\mathbf{n_0}$	**ETH**	**BTC**
100	854.82	879.59
250	637.04	698.96
500	274.08	397.93
750	0	96.89
1000	0	0

Table 1: The smallest value R_{min} of the reserves during each simulation using three years' worth of historical price data from each Ethereum (ETH) and Bitcoin (BTC), for various speculator initial endowments n_0 and initial reserves $R_0 = 1000$.

tems, we collected historical price data on Bitcoin (BTC) and Ethereum (ETH) – the two most popular cryptocurrencies in the present day – from April 2022 to April 2025. Each data point represents the hourly opening price of that currency on the online trading platform Coinbase [1], giving us a sequence of roughly 26,000 prices for each currency. We then simulated the speculator's response against each sequence to see whether reserve depletion could occur during this three-year time frame. We assumed that each of the reserves had an initial store of $R_0 = 1000$ coins, and we calculated the smallest value R_{min} the reserves hit during the course of the simulation, given different initial values for n_0. The results are shown in Table 1, where each entry in the table represents R_{min} for a particular currency, given an initial speculator endowment of n_0 coins.

Thus, we see that given a large enough initial endowment, reserve depletion eventually occurred for both types of backing coins – even when the speculator used a strategy that had not been optimized specifically for real-world prices. It is also notable that reserves of Ether depleted more quickly in practice than reserves of Bitcoin. This perhaps can be attributed to the fact that Ether historically has had higher rates of volatility than Bitcoin. To explore this hypothesis further, we examine the impact of different levels of price volatility on both the probability of reserve depletion within a certain time frame, and the time to reserve depletion for different sets of simulated prices.

In particular, the price fluctuations of stocks in financial markets are often modeled as (Gaussian) random walks (e.g., see [14,13,10], and [28,2] for a discussion specifically related to cryptocurrencies), in which the *change* in prices (rather than the prices themselves) are drawn independently at random from some normal distribution with average step size μ_{step} and standard deviation σ_{step}. In our simulations, we compute the average value of μ_{step} and σ_{step} over our three-year collection period for each BTC and ETH: The average values we computed from our datasets and used in our simulations were $\mu_{\text{step}} \approx -0.056$, $\sigma_{\text{step}} \approx 16.7$, $p_0 = \$3047.70$ for Ethereum, and $\mu_{\text{step}} \approx 1.66$, $\sigma_{\text{step}} \approx 302$, $p_0 = \$40508.04$ for Bitcoin. Then, we simulate random walks starting from the same price p_0 and with the same value of μ_{step}, but with different values of σ_{step}, to show how both the probability of reserve depletion within 100,000 time steps, as well as the expected time-to-depletion of those simulations which *do* end in reserve depletion, change with σ_{step}. For each value of σ_{step}, we run 100 different random walks and aggregate the results. The outcome of these simulations is shown in the full-text version.

For both currencies, we can see a clear trend: as the average step size σ_{step} between prices increases, so too does either the probability of reserve depletion (in the case of Bitcoin), as well as the rate of reserve depletion (for both currencies; but for Ether especially). When σ_{step} is very close to zero, the reserves are practically never depleted.

5 Future Work

Throughout the paper, we made several simplifying assumptions to either make our work tractable or to highlight what we believed to be the important features of the model without adding too much complexity. In this section, we bring attention to some of these assumptions, explaining why they were made and what impact removing them would likely have.

No competitive market. As readers may have noticed, we do not directly model a competitive market in this paper; instead, we simply argue that the price window constrains the equilibrium market price of a stablecoin. This exclusion stems from two motivations: First, we do not believe adding a market would alter the system's long-term behavior and emergent properties so long as market funds are finite: a speculator might arbitrage against the market for a time, effectively delaying reserve depletion, but once those funds are exhausted, the system's reserves face the same fate. Essentially, there is a conservation principle at play here: finite market funds and system reserves cannot withstand infinite arbitrage opportunities. The second reason we excluded a market from this model is that, though the intuitive effect is simple, we were unable to find an equally simple representation that captured the important information but did not over-complicate our model. In future work, we would like to add a minimal representation of a market, akin to our notion of a sensitive speculator.

A single speculator. In our model, we assume there are no other agents besides a single speculator. This is done primarily for the sake of simplicity, and because the underlying dynamics of a single speculator on the system are similar even in the presence of other agents, if somewhat dampened and obfuscated beneath the other signals. In particular, multiple non-strategic speculators have an additive impact on reserve depletion, and simply decrease the expected time to reserve depletion. The presence of other non-speculative agents (who wish to buy/sell stablecoins only to use them, not to profit from arbitrage over price fluctuations) can extend the time to reserve depletion by adding additional funds to the stablecoin system's reserves; but if the agents are non-speculative, then there will not be a strong correlation between buy/sell times of these agents and the price of the backing coin. Thus, the underlying impact of the speculator is still present, if somewhat obscured by other non-speculative trades. However, modeling the interactions with multiple *strategic* speculators would be an interesting extension of this work.

Secondary stabilization mechanisms. This paper showed that a mechanism which depends on price windows as its only stabilization mechanism cannot be both short- and long-term stable. In our opinion, the most interesting extension of this work would be to understand how and whether secondary stabilization mechanisms, such as governance or equity tokens, would change this behavior. Many stablecoin designs include such "dual" tokens, the idea being to sell these tokens cheaply when reserves are low to raise funds, and to allow agents to redeem them at higher prices in the future. In the spirit of our stated goal of formalizing a mathematical framework, we would like to understand not just how known secondary mechanisms impact the probability and rate of reserve depletion, but also whether the known designs represent the entirety of the possible design space. For this, we suspect that a full or partial axiomatization of the stablecoin design space could help us restrict the scope of these secondary mechanisms.

Acknowledgments

This work was supported by Input Output (iog.io) through their funding of the Edinburgh Blockchain Technology Lab (BTL) and the IO Research Hub (IORH). This work was also supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/Y003624/1. The authors would also like to thank Aggelos Kiayias for useful discussions during the initial stages of this paper.

References

1. Coinbase. `https://www.coinbase.com/`, accessed: 2025-05-28
2. Aggarwal, D.: Do bitcoins follow a random walk model? Research in Economics **73**(1), 15–22 (2019)
3. Anker-Sørensen, L., Zetzsche, D.A.: From centralized to decentralized finance: The issue of "fake-DeFi". SSRN Electron. J. (2021)
4. Brown, P.N.: Incentives for crypto-collateralized digital assets. Proceedings **28**(1) (2019). `https://doi.org/10.3390/proceedings2019028002`, `https://www.mdpi.com/2504-3900/28/1/2`
5. Bullmann, D., Klemm, J., Pinna, A.: In search for stability in crypto-assets: are stablecoins the solution? Occasional Paper Series 230, European Central Bank (Aug 2019). `https://doi.org/None`, `https://ideas.repec.org/p/ecb/ecbops/2019230.html`
6. Cao, Y., Dai, M., Kou, S., Li, L., Yang, C.: Designing Stable Coins. SSRN Electronic Journal (2021). `https://doi.org/10.2139/ssrn.3856569`
7. Catalini, C., De Gortari, A.: On the Economic Design of Stablecoins. SSRN Electronic Journal (2021). `https://doi.org/10.2139/ssrn.3899499`

8. Catalini, C., De Gortari, A., Shah, N.: Some Simple Economics of Stablecoins. SSRN Electronic Journal (2021). https://doi.org/10.2139/ssrn.3985699
9. Clements, R.: Built to fail: The inherent fragility of algorithmic stablecoins. Wake Forest L. Rev. Online **11**, 131 (2021)
10. Cootner, P.H.: The random character of stock market prices. MIT Press (1964)
11. d'Avernas, A., Maurin, V., Vandeweyer, Q.: Can Stablecoins Be Stable? SSRN Electronic Journal (2022). https://doi.org/10.2139/ssrn.4226027
12. Duan, K., Urquhart, A.: The instability of stablecoins. Finance Research Letters **52**, 103573 (Mar 2023). https://doi.org/10.1016/j.frl.2022.103573
13. Fama, E.F.: The behavior of stock-market prices. The journal of Business **38**(1), 34–105 (1965)
14. Fama, E.F.: Efficient capital markets. Journal of finance **25**(2), 383–417 (1970)
15. Grobys, K., Junttila, J., Kolari, J.W., Sapkota, N.: On the stability of stablecoins. Journal of Empirical Finance **64**, 207–223 (2021)
16. Jarno, K., Kołodziejczyk, H.: Does the Design of Stablecoins Impact Their Volatility? Journal of Risk and Financial Management **14**(2), 42 (Jan 2021). https://doi.org/10.3390/jrfm14020042
17. Klages-Mundt, A., Harz, D., Gudgeon, L., Liu, J.Y., Minca, A.: Stablecoins 2.0: Economic foundations and risk-based models. In: Proceedings of the 2nd ACM Conference on Advances in Financial Technologies (AFT). pp. 59–79 (2020)
18. Klages-Mundt, A., Minca, A.: While stability lasts: A stochastic model of noncustodial stablecoins. Mathematical Finance **32**(4), 943–981 (2022)
19. Klages-Mundt, A., Minca, A., Youngblom, R.: (In) Stability for the Blockchain: Deleveraging Spirals and Stablecoin Attacks. Cryptoeconomic Systems **1**(2) (2021)
20. Kozhan, R., Viswanath-Natraj, G.: Decentralized Stablecoins and Collateral Risk. SSRN Electronic Journal (2021). https://doi.org/10.2139/ssrn.3866975
21. Li, Y., Mayer, S.: Money creation in decentralized finance: A dynamic model of stablecoin and crypto shadow banking. Fisher College of Business Working Paper (2020-03), 030 (2022)
22. Lipton, A., Sardon, A., Schär, F., Schüpbach, C.: 11. Stablecoins, digital currency, and the future of money. Building the New Economy **30** (2020)
23. Liu, J., Makarov, I., Schoar, A.: Anatomy of a Run: The Terra Luna Crash. SSRN Electronic Journal (2023). https://doi.org/10.2139/ssrn.4416677
24. Lyons, R.K., Viswanath-Natraj, G.: What keeps stablecoins stable? Journal of International Money and Finance **131**, 102777 (2023)
25. Mita, M., Ito, K., Ohsawa, S., Tanaka, H.: What is stablecoin?: A survey on price stabilization mechanisms for decentralized payment systems. In: 2019 8th international congress on advanced applied informatics (IIAI-AAI). pp. 60–66. IEEE (2019)
26. Moin, A., Sekniqi, K., Sirer, E.G.: Sok: A classification framework for stablecoin designs. In: Financial cryptography and data security: 24th international conference, FC 2020, kota kinabalu, Malaysia, february 10–14, 2020 revised selected papers 24. pp. 174–197. Springer (2020)
27. Potter, Y., Pongmala, K., Qin, K., Klages-Mundt, A., Jovanovic, P., Parlour, C., Gervais, A., Song, D.: What Drives the (In) stability of a Stablecoin? In: Proceedings of the IEEE International Conference on Blockchain and Cryptocurrency (ICBC). pp. 316–324 (2024)
28. Urquhart, A.: The inefficiency of bitcoin. Economics Letters **148**, 80–82 (2016)
29. Zhao, W., Li, H., Yuan, Y.: Understand Volatility of Algorithmic Stablecoin: Modeling, Verification and Empirical Analysis (Jan 2021)