How to preserve the singleness of money for tokenised forms of money?

Gordon Y. Liao

Tokenised money, whether it is stablecoins or other forms of tokenised liabilities such as tokenised deposits, can be exchanged against the underlying fiat values in the primary market and often also in secondary markets. The primary market services the direct customers of the token issuer. These customers include both institutional end users that have gone through the necessary account opening process as well as resellers, or arbitrageurs, that make markets on secondary markets. The secondary markets, which could be either centralized or decentralized exchanges, service a broader range of end-users.

Recent discussions have highlighted deviations from parity in the pricing of token money and concerns with preserving the singleness of money.[1] These discussions have largely overlooked the differences between primary and secondary markets and have instead attributed the price deviations to other factors. This note aims to clarify the differences between primary and secondary markets, factors affecting pricing and liquidity of token money, and proposes solutions to preserve the singleness of money in the age of tokenised money.

Primary versus secondary market pricing

Prices of stablecoins in the secondary market can deviate from $1 in contrast to at-par redemption and issuance that occurs directly with the issuer in the primary market. Large deviations from secondary market pricing of token money have sometimes been mistaken with runs, when these are in fact reflections of the illiquidity of secondary markets. Table 1 below shows that the volume of secondary market activities is dwarfed by primary market conversion volume.

Table 1: Daily volume weighted price distribution of money token in primary and secondary market[2]

		Daily volume weighted average price of USDC in USD					
Market	**Average daily volume ($millions)**	**Min**	**1%**	**Mean**	**50%**	**99%**	**Max**
Primary	$645.5	$1	$1	$1	$1	$1	$1
Secondary	$32.5	$0.9452	$0.9991	$0.9999	$1.0000	$1.0007	$1.0109
Combined	$678.0	$0.9841	$0.9999	$1.0000	$1.0000	$1.0000	$1.0004
Combined (weekdays only)		$0.9988	$1.0000	$1.0000	$1.0000	$1.0000	$1.0002

[1] Bank for International Settlements. (2023, March 10). BIS Bulletin, No. 73. Retrieved from https://www.bis.org/publ/bisbull73.htm

[2] Replicated from https://cepr.org/voxeu/columns/payment-versus-trading-stablecoins

Note: This table shows the daily price distribution of USDC in USD and the associated volume in the primary and secondary market. The primary market volume is the sum of gross daily issuance and redemption. The secondary market volume and price are based on the conversion rate of USDC to and from USD on exchanges. The price and secondary market volume data are from Kaiko. The sample period is from 1 March 2021 to 13 March 2023, inclusive of the weekend of 11 March 2023, during which the secondary market prices of USDC temporarily deviated from parity in the aftermath of the collapse of Silicon Valley Bank.

Factors determining secondary market pricing

The crucial determining factors for the stability of secondary market pricing are:

1) Amount of float held by resellers, e.g. balance sheet capacity.
2) Speed for which round trip arbitrage can occur between secondary market and primary market.
3) Size of demand shocks in the secondary market. These demand shocks can emanate from the confidence in stablecoin asset backing.

Recent research has emphasized the significance of balance sheet capacity in understanding the variations in secondary market price deviations among stablecoins.[3] It specifically highlights the possibility of heterogeneous balance sheet capacity among resellers as a determining factor. While this theory holds true in principle, it is important to note that in practice, the differences in reseller balance sheet capacity across stablecoins are limited. This is primarily because the same set of market makers actively participate in arbitrage across multiple stablecoins.

The deviation of secondary market prices from the par value is more realistically determined by the speed of arbitrage. When prices are below parity, market makers purchase stablecoins in the secondary market and redeem from the issuer in the primary market. The opposite occurs when prices are above parity. This arbitrage process occurs with high throughput and low balance sheet utilization when there are little frictions for fiat payments between issuer and market makers. For example, certain intra-bank settlement application programmable interfaces[4] enabled market makers to complete round-trip arbitrage within a few minutes. This allows them to effectively align secondary market prices with the par value without the need for balance sheet expansion.

Additionally, interbank fiat settlements through real-time gross settlement systems like Fedwire provide relatively efficient means of settling fiat funds, enabling market makers to engage in stablecoin price arbitrage. However, the interbank settlement process introduces additional frictions, as both the market maker's bank and the stablecoin issuer's bank need to process the wire transfers promptly. These processes are often limited by manual procedures and banking hours, which can slow down the settlement and arbitrage activities.

[3] Ma, Yiming, Zeng, Yao, and Zhang, Anthony Lee, 2023 "Stablecoin Runs and the Centralization of Arbitrage."
[4] Such as Silvergate Exchange Network and Signature's Signet.

Certain stablecoins utilized a batched issuance and redemption process and imposed redemption fees. This alternate model is not conducive to high-throughput arbitrage between primary and secondary markets. Consequently, these stablecoins typically exhibit greater price deviations in the secondary market. The implementation of this batched redemption process might have been required due to the lack of banking partners with direct access to the Fedwire system. The dependence on a lengthy chain of correspondent banks inevitably prolongs fiat settlement time, thus making batching a necessary procedure.

Lastly, the stability of secondary market pricing is fundamentally dependent on the trust that holders of tokenized money place in the underlying assets. If these assets are not high-quality, highly liquid or if they are fraught with significant market or credit risks, the token money could be prone to runs. Token holders might rush to redeem their tokens at the first signs of concerns with the backing asset. Programmable money also enables programmable runs. Insufficiently backed token money would also provide fertile ground for speculative attacks.

How can singleness of money be preserved for tokenised money?

Policies can be devised to maintain the singleness of money through tackling some of the factors discussed above. First, requiring a high-level of asset quality and liquidity in reserve assets for token money is paramount in preventing runs. Given the highly programmable, transferable nature of token money, these asset backing requirements must exceed those that are typical of traditional depository institutions.

Second, access to central bank money and settlement systems such as Fedwires is necessary to facilitate efficient functioning of the secondary market. As discussed above, fast fiat settlement processes are critical for liquidity provision in the secondary market. These fiat settlements should occur interbank rather than be confined to intrabank, which can create concentration risks. Allowing direct access to fiat payment rails and the ability to maintain small balances at the central bank for facilitating fiat transactions would substantially improve secondary market liquidity and reduce credit exposure to intermediaries.