Journal of Futures Markets

WILEY

RESEARCH ARTICLE OPEN ACCESS

The Economics of Liquid Staking Derivatives: Basis Determinants and Price Discovery

Stefan Scharnowski[1] | Hossein Jahanshahloo[2]

[1]Finance Area, University of Mannheim, Mannheim, Germany | [2]Department of Accounting and Finance, Cardiff Business School, Cardiff University, Cardiff, UK

Correspondence: Stefan Scharnowski (scharnowski@uni-mannheim.de)

Received: 17 February 2023 | **Revised:** 28 August 2024 | **Accepted:** 4 November 2024

Keywords: cryptocurrency | liquidity | price discovery | proof-of-stake | staking

ABSTRACT

This paper provides a first economic analysis of liquid staking tokens, which are derivatives representing a share of staked tokens in Proof-of-Stake blockchains. We document substantial time-variation in the "liquid staking basis" as given by the price difference between a derivative staking token and its underlying cryptocurrency. We find evidence that staking rewards, concentration risks, limits to arbitrage, and behavioral factors influence this basis. The liquid staking basis is wider when the yields offered by the liquid staking protocol are low relative to the alternative of staking directly, when cryptocurrency returns are more volatile, and when secondary market liquidity is low. In contrast, it is smaller when investors pay more attention to liquid staking and when investor sentiment is positive. Furthermore, liquid staking tokens contribute a significant and overall growing amount to price discovery in the underlying cryptocurrencies.

JEL Classification: G12, G13, G15, G23

1 | Introduction

This paper provides a first empirical investigation of the economics of liquid staking tokens, which are derivatives representing a share of staked tokens in Proof-of-Stake (PoS) blockchains. The PoS mechanism has become a popular alternative to the energy-intensive Proof-of-Work (PoW) mechanism. While the latter is used by Bitcoin and other cryptocurrencies, the market share of PoS blockchains such as Ethereum and Solana has increased substantially over the recent past. In these blockchains, the integrity of new transactions is verified by network participants called validators that append the transactions to the blockchain and are rewarded in the form of newly minted coins or transaction fees. The validators are incentivized to behave honestly by a requirement to have financial stake in the network which they stand to lose should they break the protocol's rules. By design, this reduces liquidity as the posted stake has to be locked up for a specific period, which makes it unavailable for other applications.

In direct staking, validators deposit the stake directly within the protocol and hence suffer from the reduced liquidity. Liquid staking providers address this by introducing derivative tokens that represent a share of a pool of staked assets. Liquid staking as an alternative to direct staking thus provides a layer of intermediation, offering several advantages. First, the liquid staking tokens can be traded or used as collateral in decentralized lending platforms. Second, liquid staking tokens can be sold in the secondary market, sometimes without any lock-up period. Third, liquid staking typically does not have minimum stake requirements and thus lower barriers to entry. However, there are some drawbacks compared to direct staking. Liquid staking providers keep parts of the rewards, thus decreasing staking yields and increasing yield volatility. Additionally, there are concentration risks if a single provider controls a significant share of the staked amount. Furthermore, the staking platforms may be vulnerable to attacks.

The market capitalization and trading activity of liquid staking tokens have recently seen substantial growth. For example, the

https://doi.org/10.1002/fut.22556

market capitalization of the largest liquid staking token at the time of writing, Lido Staked Ether, has grown from about USD 20 million in January 2021 to USD 15 billion in July 2023, accounting for more than 30% of all deposited stake in Ethereum. However, even though the market for liquid staking tokens has grown significantly, little is known about the tokens' price dynamics.

In this study, we attempt to fill this gap by examining the tokens' pricing and their contribution to the price discovery process of the underlying cryptocurrencies. We start our investigation by studying price differences between the derivative liquid staking tokens and the underlying cryptocurrencies, referred to as the “liquid staking basis”. Building on the literature on the pricing of other derivatives, we then look at the potential determinants of this basis. We find that four aspects are particularly relevant for the relative pricing of liquid staking tokens: Staking rewards, concentration risks, limits to arbitrage, and behavioral factors. For staking rewards, we find that the basis is wider when liquid staking offers relatively low yields compared the alternative of staking directly. We also find that an increase in concentration risk is associated with a wider basis, suggesting that traders are aware of this risk. Regarding limits to arbitrage, our findings suggest that the basis is inversely related to the secondary market liquidity of staking tokens, consistent with the notion that low liquidity hinders the activity of arbitrageurs. In particular, the basis increases when decentralized exchange fees or centralized exchange bid-ask spreads increase. Similarly, funding illiquidity as measured by forced liquidations due to margin calls is positively related to the magnitude of the basis. For the behavioral factors, we find that the basis is smaller when investors pay more attention to liquid staking. Additionally, market sentiment plays an important role as the basis is substantially wider when the implied volatility of other cryptocurrency derivatives is higher.

While liquid staking tokens have a different economic function than cryptocurrency futures contracts, their pricing in relation to the underlying asset may exhibit similarities. To study this aspect, we compare the liquid staking basis to the futures-spot basis. We find that there is some overlap between the determinants of the two bases, although they are not identical due to differences in market structure and maturity.

Finally, we consider how the liquid staking tokens contribute to the price discovery process of the underlying cryptocurrencies. Using information shares, we find that – compared to the spot and futures market – these derivative tokens contribute a significant and growing amount to price discovery, which suggests an increase in importance of the tokens within the DeFi ecosystem.

Our findings bear significant implications for traders and DeFi market participants. A comprehensive understanding of the pricing dynamics of these derivative tokens is of paramount importance for effective risk management. Likewise, our findings are relevant for financial regulators aiming at efficiently regulating cryptocurrencies and related assets, especially since liquid staking tokens played a prominent role in the turmoil in cryptocurrency markets starting May 2022 and may thus have implications for systemic risk within decentralized finance.

We contribute to three related streams within the literature. Firstly, we contribute to the literature on the economics of different blockchain consensus mechanisms by empirically examining a potentially important mechanism for adding liquidity to PoS networks. Within this stream, Saleh (2021) and Roşu and Saleh (2021) investigate the equilibrium characteristics of the PoS consensus mechanism. Cong, He, and Tang (2022), Choi, Jeon, and Lim (2023), and John, Rivera, and Saleh (2021) formalize the trade-off between holding tokens in a liquid account, thus earning convenience yields for example by using them for economic transactions, and staking the tokens to earn rewards. Sapkota and Grobys (2021) document that PoS cryptocurrencies on average have similar returns to those utilizing the PoW consensus mechanism. Milunovich (2022) looks at linkages between cryptocurrencies using different consensus mechanisms and finds that PoS currencies on average are not as strongly connected to other cryptocurrencies than those using PoW. Explicitly considering liquid staking, Gogol et al. (2024) classify liquid staking providers and analyze the performance of major liquid staking derivatives. Lastly, Tzinas and Zindros (2024) investigate the principal-agent conflicts that can emerge in liquid staking due to the liquid representation of the pooled delegation of stakes to validators. They outline potential attacks that exploit this problem and suggest ways to make liquid staking systems more robust.

Secondly, we contribute to the literature on cryptocurrency derivatives, which has mostly been focused on futures contracts.[1] Looking at price differences between Bitcoin futures and spot markets, Hattori and Ishida (2021) show that these potential arbitrage opportunities are rare during calm periods, but more common during market crashes. Schmeling, Schrimpf, and Todorov (2023) examine the futures-spot basis in cryptocurrency markets, called *crypto carry*, and show that it is volatile and large in magnitude. They also identify the behavior of retail investors and scarce arbitrage capital as important determinants of this basis. Several studies investigate the contribution of the futures market to the price discovery process. While Corbet et al. (2018) and Baur and Dimpfl (2019) find that the spot market leads price discovery of the underlying asset, subsequent studies conclude that futures markets contribute significantly to the price discovery process (Kapar and Olmo 2019; Aleti and Mizrach 2021; Hu, Hou, and Oxley 2020; Alexander et al. 2020a; Alexander and Heck 2020). Entrop, Frijns, and Seruset (2020) document time-variation in the contribution of the futures market to the price discovery of Bitcoin and identify trading costs, volume, sentiment, uncertainty, and the size of a trade to be relevant determinants of price discovery. De Blasis and Webb (2022) compare the market structure of quarterly and perpetual futures contracts, which are mostly unique to cryptocurrency markets, and find return spillovers between both types of contracts as well as arbitrage opportunities. We contribute to this stream of the literature by investigating a new type of liquid cryptocurrency derivative and its contribution to the price discovery process of the underlying cryptocurrencies.

Thirdly and more generally, we contribute to the literature on decentralized finance. Important components of DeFi such as stablecoins (Griffin and Shams 2020; Hoang and Baur 2021; Grobys et al. 2021), decentralized exchanges (Lehar and

Parlour 2021; Aspris et al. 2021; Lo and Medda 2021), and decentralized lending platforms (Chiu et al. 2023; Lehar and Parlour 2022) have already been investigated in the literature. For many DeFi market participants, liquid staking tokens now form an essential part of their strategies, for example by using the staking tokens as collateral when taking out cryptocurrency loans. Understanding their price dynamics is hence crucial for DeFi market participants.

The rest of the paper is organized as follows. Section 2 provides the technological and institutional background of liquid staking. Section 3 develops the hypotheses. Section 4 introduces the data and the empirical methodology. Section 5 shows the development of the liquid staking basis and investigates its potential determinants. Section 6 illustrates the staking tokens' contribution to price discovery and Section 7 concludes.

2 | Institutional Background

2.1 | Proof-of-Stake and Liquid Staking

Cryptocurrencies and other digital assets have increased significantly in both popularity and market capitalization since they first emerged. There are now thousands of different such digital assets available. One of the main properties that distinguishes different decentralized cryptocurrencies is how the respective network agrees on the correct version of the distributed ledger that keeps track of all transactions or balances. Some networks, such as Bitcoin, use the PoW consensus mechanism where miners solve complex but otherwise meaningless mathematical puzzles. The most popular alternative to the energy-intensive PoW mechanism is PoS, used by cryptocurrencies such as Solana, Cardano, and, more recently, Ethereum. (Irresberger et al. 2020).

In PoS, transactions are verified by randomly chosen validators. To participate, validators have to deposit a number of coins as stake, which are locked up for some time to incentivize the validators to behave honestly since they stand to lose the locked value should they compromise the blockchain. The probability of being chosen as validator generally increases in the amount deposited as stake.[2] In return, the validators receive some reward. Depending on the design of the blockchain, this reward typically contains two components. Firstly, validators receive the relatively stable consensus layer reward for attesting transactions in the form of newly minted coins. Moreover, rewards include the more volatile execution layer reward in the form of transaction fees and maximal extractable value (MEV), i.e., the profits that can be extracted from a proposed block by ordering and sequencing the transactions.

By design, the process of locking up the stake reduces liquidity, since the amount cannot be used for any other purpose (Cong, He, and Tang 2022; Choi, Jeon, and Lim 2023). Moreover, there are technological and financial barriers for network participants to participate in the validation process directly. While contrary to PoW, becoming a validator does not require highly specialized hardware, it still requires detailed technical knowledge. Furthermore, in some cases a minimum amount has to be staked which may be prohibitively high for smaller network participants. For example, to become a validator in Ethereum, a stake of at least 32 Ether (ETH) is required (Grandjean, Heimbach, and Wattenhofer 2023).

To address these issues, there are several alternatives to direct staking.[3] In staking-as-a-service, a user deposits her stake with a third party that runs and maintains the staking infrastructure for a fee. This process requires a high level of trust in the third party and does not address the liquidity reduction stemming from the locked-up stake. Similarly, several centralized cryptocurrency exchanges allow users to stake deposited tokens.[4] The exchange then operates its own validation infrastructure or forwards the stake to third parties. Correspondingly, this process is usually custodial, requiring users to hand over control of their funds to the exchange as in staking-as-a-service. In pooled staking, several users pool their stake and validate transactions without managing individual validator infrastructures. To this end, pools may operate their own staking infrastructure or distribute the stake among one or multiple third parties operating such infrastructures. Pooled staking can be custodial or non-custodial. In the former, users have to hand over control of the stake to the pool operator, similar to staking with a centralized exchange. The latter is facilitated by smart contracts where the pool operator cannot access the funds of the pool participants.

In this paper, we investigate a popular form of non-custodial pooled staking: liquid staking. Interested parties transfer their coins to the smart contracts of these services which then pool the stakes and distribute them among a set of validators. In return, users of these services receive derivative tokens that represent their share in the staking pool. As investors stake and unstake their positions, the liquid staking provider mints and burns (i.e., creates and deletes) these liquid staking tokens. Liquid staking providers can either be centralized entities or decentralized autonomous organizations (DAOs), each with different governance mechanisms.[5]

The incurred staking rewards, both in terms of the block rewards for successfully validating a block and in terms of the fees paid by users to have their transactions included in a block, are used to cover the expenses of the service providers and of the validators. The remainder is paid out to those providing the stake. The value of the positions of those holding liquid staking tokens hence increases over time, although the exact payout mechanics differ between liquid staking protocols and generally fall into three categories. Some protocols use rebasable tokens where rewards are paid out by increasing the total supply of liquid staking tokens outstanding and proportionally distributing these newly issued tokens among the staking participants. The prices of rebasable tokens hence do not increase due to the rewards; instead, the token balances of liquid staking participants increase over time. While intuitive, the downside of this approach is a reduced compatibility with other DeFi applications as not all support rebasable tokens. In contrast, for reward-bearing liquid staking tokens, each token reflects the same share of a pool that increases in value as rewards are added to the pool. An individual token hence appreciates in value over time, mechanically increasing its price in the secondary market. Some protocols offer both rebasable and reward-bearing tokens which can be exchanged

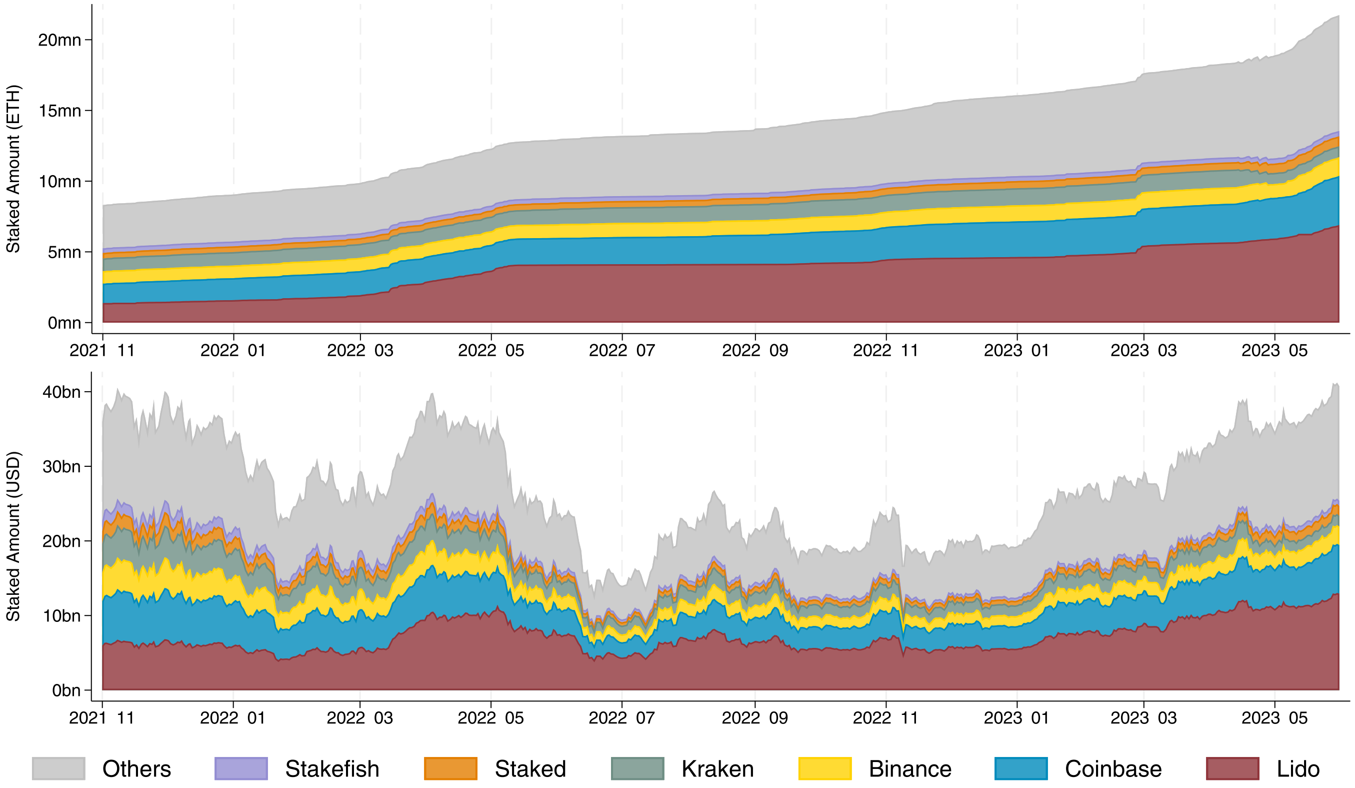


FIGURE 1 | Staked amount in Ethereum. *Note:* These graphs show the total staked amount in Ethereum. The six largest staking services during the sample period are shown individually. *Others* includes all other forms of staking, including direct staking. The individual areas are stacked. The staked amount is given in Ether and USD in the top and bottom graph, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

via a smart contract in a process called "wrapping" and "unwrapping". Finally, a less common payout mechanic is that of a dual token model where rewards are paid out in the form of a separate token. Liquid staking participants thus hold two types of tokens, one reflecting their share of the stake and one for the rewards.

In general, liquid staking tokens can increase liquidity in two ways: Firstly, the derivative tokens can themselves be traded on centralized or decentralized exchanges or used in DeFi applications, for example as collateral for cryptocurrency loans on decentralized lending platforms. Liquid staking tokens hence allow market participants to obtain some convenience yields while simultaneously earning staking rewards, reducing the typical trade-off between the two as discussed for example in Cong, He, and Tang (2022) and Choi, Jeon, and Lim (2023). Secondly, some liquid staking services allow token holders to immediately (i.e., without any lock-up period) withdraw their stake for a fee.

However, liquid staking also comes with risks. In particular, liquid staking adds another layer of intermediation which introduces smart contract risk, as the liquid staking protocol itself may be vulnerable to attacks. Furthermore, the liquid staking tokens and the venues where these can be traded potentially suffer from liquidity and counterparty risk while the yields – or rather the yield difference to staking the underlying asset directly – are volatile. Investors hence have to trade-off these risks and the benefits of the added liquidity.

2.2 | The Market for Liquid Staking Derivatives

Two of the largest liquid staking service providers are Lido and Marinade.[6] Lido offers staking services for various cryptocurrencies, including Ethereum, Solana, Polygon, and Polkadot, whereas Marinade specializes in offering services for Solana. At the time of writing, the two largest markets for liquid staking tokens are Ethereum and Solana.

The market for liquid staking tokens has grown substantially over the recent past. The aforementioned example, Lido Staked Ether (stETH), has grown from about USD 20 million in January 2021 to USD 15 billion in July 2023. However, the importance of liquid staking tokens has increased not only in absolute terms, but also relative to the overall amounts staked on the various blockchains. Figure 1 illustrates the amount staked during the sample period of our study. Staking services, including both custodial and non-custodial services, constitute the majority of staked Ether. Importantly, the liquid staking provider Lido accounts for more than 30% of all staked Ether, further highlighting the relevance of a thorough understanding of the economics of this market. The high market share of a few providers also raises centralization and governance concerns (see e.g. Ethereum Foundation 2022; Grandjean, Heimbach, and Wattenhofer 2023).

During our sample period, Ethereum transitioned from using the PoW consensus mechanism to PoS. However, the staking mechanics were already implemented in the form of the

"Beacon chain", Ethereum's PoS consensus layer. Since late 2020, this layer was running in parallel to the PoW mechanism, allowing users to participate in staking to earn rewards. On September 15th, 2022, the PoS consensus layer was merged with the Mainnet that keeps track of all transactions and the global state of the network. Starting on this date, all transactions are validated by PoS validators instead of PoW miners. The Merge also has important implications for staking rewards. While before, validators only received the consensus layer rewards in the form of newly minted Ether, afterwards they additionally receive the more volatile execution layer rewards in the form of transaction fees and MEV.

However, neither before nor immediately after this merge it was possible to un-stake in Ethereum, a feature that was added to the network as part of the "Capella" and "Shanghai" hard forks. Both of these upgrades to the Ethereum network took place on April 12th, 2023.[7] For those staking in Ethereum via a liquid staking token, exiting a staking position by redeeming the staking tokens for the underlying stake is hence only possible after the transition completed. In the case of Lido, the functionality to withdraw staked ETH from the protocol was added in May 2023. In the meantime, users could only exit their positions by selling them in the secondary market. Conversely, staking is fully implemented in Solana during our sample period and users can already redeem their staking tokens for the underlying stake and the accrued rewards.

Liquid staking tokens are widely used throughout various DeFi applications. For example, they can be traded in secondary markets including decentralized exchanges (DEXs) such as Curve. Token holders can also earn yields by supplying liquidity to DEXs. For example, a holder of both stETH and ETH could supply both assets to the stETH/ETH liquidity pool of a DEX, thus earning staking rewards on the stETH position and additional rewards from providing liquidity to the DEX.

Liquid staking tokens can also be used as collateral on lending platforms. A popular trading strategy involving these tokens is given by first staking the native token, e.g., ETH, via a liquid staking protocol. The corresponding liquid staking token, e.g., stETH, is then used as collateral on a decentralized lending platform such as Aave to borrow the native token, which is then staked again. This process can be repeated several times depending on the degree of (over-)collateralization on the lending platform. This recursive strategy, sometimes referred to as "folding", allows traders to enter leveraged staking positions (Coinbase Institutional 2023).

Strategies involving liquid staking tokens economically link different platforms and increase the tokens' importance within DeFi. While this is a desired consequence of the added liquidity of liquid staking, it also increases systemic risks. This became apparent during the widespread turmoil in cryptocurrency markets in 2022, commonly referred to as the "crypto winter". On May 9th, 2022, the stablecoin TerraUSD (UST) broke its peg to the US dollar, subsequently dropping to a price of almost zero while losing about $45 billion in market capitalization. This event triggered a market-wide downturn, leading to the bankruptcies of large cryptocurrency institutions such as Three Arrows Capital, Voyager Digital, and Celsius Network. The latter was one of the largest holders of stETH and had pledged about USD 400 million of the token on the lending platform Aave (OECD 2022). Following a de-pegging of stETH and ETH, partially caused by large sales of stETH by Alameda Research, the trading arm of the centralized exchange FTX, Celsius and other institutions were forced to liquidate positions in stETH. This not only led to further imbalances in secondary markets for liquid staking tokens, but also to spillovers to other markets (OECD 2022).

3 | Hypothesis Development

There are several reasons to expect that the liquid staking basis, i.e., the price difference between a liquid staking token and its underlying asset, is non-zero and varies over time. In this section, we develop four testable hypotheses regarding its potential determinants to gain a better understanding of the relative pricing of these derivatives. The hypotheses relate to the incentives of staking, concentration risks, limits to arbitrage, and behavioral factors.

3.1 | Staking Incentives

One of the main incentives to participate in staking is to receive rewards in the form of new tokens, fees, or both (Saleh 2021; Cong, He, and Tang 2022). This reward is in itself volatile, in particular since the execution layer rewards, i.e., fees and MEV, vary substantially over time. This volatility impacts all types of staking, including direct and liquid staking. However, liquid staking providers keep parts of these rewards to cover their expenses and to compensate the validators. This share also varies over time. Investors in liquid staking derivatives hence forego some of the rewards of staking directly. Economically, the effect is similar to how dividend payments affect the relative pricing of equity futures contracts (Cornell and French 1983). Analogously, we expect liquid staking to be particularly attractive if the difference in yields from staking the underlying asset directly and from liquid staking is relatively small. Conversely, if this yield difference is large, liquid staking becomes less attractive, decreasing the prices of liquid staking tokens relative to the underlying cryptocurrencies. Following this reasoning, we conjecture that:

Hypothesis 1. *The liquid staking basis increases when the yield difference between staking directly and liquid staking increases.*

Before the Merge, validators only received the consensus layer rewards in the form of newly minted coins. Afterwards, they additionally receive the execution layer rewards in the form of transaction fees and MEV which is usually more volatile. Because the reward structure changes after this event, we also contrast the effect of yield differences on the basis before and after the Merge.

3.2 | Concentration Risk

For Lido Staked Ether in particular, market participants are concerned about the concentration of the stake deposited in the protocol since high levels of centralization potentially make the entire network more vulnerable to attacks (Grandjean, Heimbach, and Wattenhofer 2023). This concentration risk may be reflected in the liquid staking basis as investors require a premium for bearing this risk. The basis would then increase when Lido's share of overall stake increases. We hence hypothesize that:

Hypothesis 2. *The liquid staking basis increases when Lido's share of staking increases.*

3.3 | Limits to Arbitrage

Price differences between a derivative and its underlying asset are generally larger if market conditions make exploiting these potential arbitrage opportunities difficult. For example, high price volatility increases the risks faced by arbitrageurs, thus limiting their ability to conduct arbitrage trading (Shleifer and Vishny 1997; Gagnon and Karolyi 2010). Furthermore, there may be asymmetries in how arbitrageurs react to increasing or decreasing prices, for example due to funding or short selling constraints (Stambaugh, Yu, and Yuan 2015).

As discussed by Roll, Schwartz, and Subrahmanyam (2007) and further examined by Kadapakkam and Kumar (2013) and Han and Pan (2017), deviations from the law of one price can be partially attributed to illiquidity, since arbitrage becomes more costly when liquidity is low. Although the focus of the studies above is on the equity index futures basis, the same logic applies to the liquid staking basis. It is crucial in our context to differentiate between various types of liquidity. Liquid staking tokens enhance liquidity for those staking their assets by effectively securitizing their stake. However, the liquidity of these tokens in the secondary market may vary over time or between different staking tokens, resulting in temporal variations in potential arbitrage opportunities.

Additionally, a lack of trading activity in staking tokens can make arbitrage more challenging. However, price differences may also arise from imbalances, such as one-sided trading activity. Moreover, high trading volume might indicate noise trader risk, which can deter potential arbitrageurs (De Long et al. 1990), leading to a wider basis.

In a similar vein, there may be spillovers from funding liquidity to market liquidity as in Brunnermeier and Pedersen (2009). During times of low funding liquidity, capital-constrained investors may need to quickly liquidate their positions. If they sell their liquid staking derivatives first, which may simply be more accessible compared to locked-up stake, the tokens' prices would decrease relative to the underlying asset. Likewise, funding illiquidity may make arbitrage activity more difficult and expensive (see also Schmeling, Schrimpf, and Todorov 2023).

Overall, high volatility, low liquidity, and funding constraints hinder the exploitation of potential arbitrage opportunities. These limits to arbitrage might increase the price difference between liquid staking tokens and their underlying cryptocurrencies. Accordingly, we hypothesize:

Hypothesis 3. *Limits to arbitrage such as high price volatility and low liquidity are positively related to the liquid staking basis.*

3.4 | Investor Behavior

Price differences may be more pervasive if investors are not aware of – or paying particular attention to – liquid staking tokens, thus missing out on the opportunity to exploit existing mispricing. Eichler (2012) considers mispricing in American Depositary Receipts and finds that price differences to the underlying stock decrease with an increase in investor attention. Similarly, in a study on commodity futures markets, Han, Li, and Yin (2017) find that increased investor attention, as measured by Google search volume, is associated with fewer arbitrage opportunities. We hence hypothesize that the liquid staking basis is lower in magnitude if investor attention is high.

In addition to attention, investor sentiment is also known to affect asset prices and potentially drive these away from their fundamentals, especially for assets that are more difficult to arbitrage (Baker and Wurgler 2006; Han et al. 2022). In particular, during times of low investor sentiment, risk appetite and risk absorption capacities of potential arbitrageurs may be lower (Cheng, Kirilenko, and Xiong 2015).

We hence expect that both a lack of investor attention and negative investor sentiment regarding the cryptocurrency market is associated with a wider liquid staking basis.

Hypothesis 4. *The liquid staking basis is inversely related to investor attention and sentiment.*

4 | Data and Empirical Methodology

4.1 | Data

We obtain high frequency trading data from two markets for liquid staking tokens: Curve, a DEX, and FTX, a centralized cryptocurrency exchange (CEX). We consider these two markets for several reasons. The DEX data captures a significantly larger fraction of overall trading activity in liquid staking tokens. However, the CEX data allows us to compare trading activity in different assets at the exact same trading venue. Moreover, some variables used for testing our hypotheses are only available for one of the two markets. Finally, we consider the CEX analysis a robustness test. Even though the mechanics of trading at a DEX and a CEX differ greatly (Barbon and Ranaldo 2023), testing our hypotheses in both types of markets allows us to better understand the underlying economics of liquid staking tokens, for example with respect to secondary market liquidity.

Our main analysis is based on data from Curve. From the Ethereum blockchain, we collect all transactions ("swaps") from the liquidity pool where Lido Staked Ether (stETH) can be traded against Ether (ETH). At the time of writing, Lido Staked Ether is by far the largest liquid staking token and the Curve liquidity pool the largest venue for trading this asset.[8] The sample ranges from November 1th, 2021, to May 31th, 2023. We additionally collect data on pool liquidity flows, fees, and liquidity utilization from IntoTheBlock (2023).

To supplement this analysis, we further collect trading data from FTX (2022). Although now defunct, at the time it was one of the largest and most liquid CEXs for liquid staking tokens.[9] As in our main analysis, we collect data for Lido Staked Ether (stETH), allowing us to compare the two trading venues. However, we additionally collect data on Marinade Staked Solana (mSOL) and Lido Staked Solana (stSOL), which are substantially less liquid than stETH, but during our sample among the largest liquid staking tokens by market capitalization. The raw data is sampled in 15-second intervals.[10] We end the sample period for this market on August 31th, 2022, well before the problems that eventually led to the downfall of FTX and its trading arm, Alameda Research, became publicly known. This makes it unlikely that the choice of trading venue impacts our results. Furthermore, we collect price data on the underlying cryptocurrencies Ether (ETH) and Solana (SOL), and for the two corresponding perpetual futures contracts (which we denote fETH and fSOL, respectively). We additionally collect data on forced liquidations, i.e., automatically triggered transactions due to margin calls, in all of these assets. All prices in the FTX sample are given in USD.

Since for some of our DEX-based analyses, we need cryptocurrency spot exchange rates against the USD for the period after the collapse of FTX, we collect transaction prices from Kraken (2023), a prominent and highly liquid CEX. We also collect price data on the Deribit (2023) volatility index (DVOL), which measures the implied volatility of cryptocurrency options and is thus a forward-looking measure of expected volatility which we use as a proxy for investor sentiment. Finally, we collect data on Ethereum staking rewards, the total amount of Ether staked, and the amount of Ether staked in various protocols such as Lido from Dune Analytics (2023).

Unless otherwise indicated, for our analyses we aggregate the data to 1-h intervals to avoid microstructure noise. In particular, all regressions and summary statistics are based on hourly data. However, our analysis of price discovery relies on data at the 15 second frequency.

4.2 | Measuring the Liquid Staking Basis

The liquid staking basis is defined as the relative price difference between the liquid staking tokens and their respective underlying cryptocurrency. In most of our analyses, we consider the absolute value of this basis to capture the magnitude of price differences. The calculation of the basis depends on the trading venue. For the DEX, we calculate the basis based on swaps of the liquid staking token and the underlying cryptocurrency.

$$\text{Liquid Staking Basis}^{\text{DEX}}_{\text{signed}} = \frac{\text{Q}_{\text{Underlying}}}{\text{Q}_{\text{LST}}} - 1 \tag{1}$$

$$\text{Liquid Staking Basis}^{\text{DEX}} = \left|\text{Liquid Staking Basis}^{\text{DEX}}_{\text{signed}}\right| \tag{2}$$

where Q_{LST} and $Q_{\text{Underlying}}$ denote the quantity of the liquid staking token and the underlying cryptocurrency exchanged in a swap, respectively. For the CEX, all prices are expressed in USD. The basis is hence given by

$$\text{Liquid Staking Basis}^{\text{CEX}}_{\text{signed}} = \frac{\text{Close}_{\text{LST}} - \text{Close}_{\text{Underlying}}}{\text{Close}_{\text{Underlying}}} \tag{3}$$

$$\text{Liquid Staking Basis}^{\text{CEX}} = \left|\text{Liquid Staking Basis}^{\text{CEX}}_{\text{signed}}\right| \tag{4}$$

The tokens in our sample differ with respect to the reward payout mechanics, with important implications for the calculation of the basis. In particular, stETH is a rebasing token, meaning that rewards are paid out in the form of newly issued tokens which are proportionally distributed among the staking participants, hence increasing the supply of outstanding liquid staking tokens. The price of each token is then not directly impacted by the payout. In contrast, both the Solana staking tokens traded at the CEX, stSOL and mSOL, are reward-bearing tokens. Each token thus reflects the same share of a pool that increases in value. An individual token hence appreciates in value over time.[11] To address this issue in our analysis, we make the staking tokens comparable by de-trending the liquid staking basis for these staking tokens. We de-trend the time-series of the liquid staking basis by running a first-stage regression of the basis on a linear time trend for each Solana-based staking token individually. Then, we use the residuals of these regressions in our analyses. However, to illustrate the mechanism, we present summary statistics and figures of the basis before this adjustment.

4.3 | Regression Approach and Variable Construction

Our main analysis of the determinants of the liquid staking basis consists of fixed effects regressions, first using the DEX data and then using the CEX data. We take the absolute basis as given by Equation (2) for the DEX and Equation (4) for the CEX as dependent variables, respectively. The analysis of the CEX differs from the one for the DEX in other respects due to differences in variable construction in three main areas. First, market liquidity is measured differently. Second, using the CEX data, we can more directly measure funding liquidity and test the relationship between liquid staking tokens and futures contracts trading at the same venue. Third, since the collapse of FTX occurred before the Merge, this specification does not include variables related to this event.[12]

Specifically, we estimate the following regression model:

$$\begin{aligned} \text{Basis}_t = \alpha_0 &+ \beta_1 \text{Staking Incentives}_t \\ &+ \beta_2 \text{Concentration Risk}_t \\ &+ \beta_3 \text{Limits to Arbitrage}_t \\ &+ \beta_4 \text{Investor Behavior}_t \\ &+ \beta_5 \text{Controls}_t + \gamma_1 \text{Hour}_t + \varepsilon_t \end{aligned} \tag{5}$$

where $Basis_t$ is the absolute liquid staking basis at time t.

To evaluate Hypothesis 1, *Staking Incentives* is a vector containing variables capturing incentives for participating in liquid staking. For both exchanges, it includes Δ *Staking Rewards* as given by the difference between the current annualized percentage yield of staking Ether directly and staking via Lido. For the DEX, we interact this variable with *Post Merge*, an indicator variable for the period after the Merge to capture differences in this effect after the uncertainty reduction and the change in reward structure following the Merge.

As discussed in Hypothesis 2, market participants are concerned regarding *Concentration Risks* with respect to the stake deposited in the protocol, particularly for Lido Staked Ether. To understand how this concentration risk may affect the basis, we include *Lido Share of Staking* as given by the share of Ether staked via Lido relative to the total staked amount in the Ethereum network.

Limits to Arbitrage is a vector capturing volatility and liquidity as discussed in Hypothesis 3. To proxy for volatility at the DEX and the CEX, it contains $Return^{Absolute}_{Underlying}$ as given by the absolute logarithmic return of the underlying asset. Inspired by Chordia, Roll, and Subrahmanyam (2001), the vector furthermore includes a binary variable indicating if these returns are positive to capture any asymmetries in how the basis reacts to price increases and decreases. The vector furthermore includes various liquidity measures. For the DEX, it comprises the following: *Pool Size*, represented by the logarithm of the total USD value of the stETH/ETH liquidity pool; *Trading Volume*, indicated by the USD value exchanged within the pool; *Pool Fees*, calculated as the fees paid to liquidity providers and expressed as annualized percentage yields; and *Net Liquidity Flow*, defined as the net sum of the pool's inflows and outflows. For the CEX, it comprises: $Spread_{StakingToken}$, estimated using the Corwin and Schultz (2012) high-low spread estimator as suggested for cryptocurrency markets by Brauneis et al. (2021)[13]; $Volume_{StakingToken}$ as measured by the total traded volume at FTX in the staking token in USD; and *Forced Liquidations*, serving as a proxy for funding liquidity, calculated as the daily fraction of trading volume in all assets related to a given underlying cryptocurrency (i.e., staking derivatives, futures, and spot market) that results from the automatic unwinding of traders' positions due to margin calls. A higher value of *Forced Liquidations* indicates that more traders are liquidity-constrained.

Investor Behavior is a vector that captures the influence of behavioral factors, namely investor attention and sentiment, as outlined in Hypothesis 4. As is common in the literature on investor attention (e.g. Han, Li, and Yin 2017), we proxy for *Attention* by obtaining data on Google search intensity for the term “liquid staking”. While there is no consensus in the literature on how to best measure investor sentiment, this issue is complicated by the fact that our analysis requires a high-frequency measure. A commonly used proxy in this setting can be found in implied volatility indices such as the VIX, which measure “investor fear” (Kurov 2010). Following this approach, we include *Sentiment* as measured by the level of the Deribit Volatility Index (DVOL), which tracks the implied volatility of cryptocurrency derivative contracts.

We additionally include the vector *Controls* containing several variables. For both DEX and CEX, it consists of variables capturing the crypto winter, the weekend effect, and the sign of the liquid staking basis. For the DEX, we additionally control for the Merge.

First, we control for the crypto winter of 2022. This period impacted all cryptocurrencies, including liquid staking tokens (Nansen 2022; OECD 2022), as some investors were forced to quickly liquidate their liquid staking positions by selling them in the secondary market (Nansen 2022). We hence include the binary variable *Crypto Winter* indicating the period starting May 9th, 2022, when TerraUSD broke its peg to the US dollar, until the end of the sample. In addition to the TerraUSD collapse, this period also includes the downfalls of Three Arrows Capital, Voyager Digital, Celsius Network, and FTX.

Second, we account for the weekend effect since the activity of investors and the composition of investor types may vary over time. For example, Jahanshahloo, Corbet, and Oxley (2022) show that Bitcoin on-chain activity is different on the weekend than on weekdays. Examining patterns in Bitcoin liquidity, Scharnowski (2021) provides evidence that institutional investors are generally more active during the week. If institutional investors are better at exploiting mispricing, there could be a difference in the basis throughout the week relative to the weekend. Furthermore, the basis may also differ between weekdays and the weekend due to changes in risk (see e.g. Singal and Tayal 2020). To account for these possible differences throughout the week, we include a binary variable for the *Weekend* which is equal to 1 on Saturdays and Sundays[14].

Third, there may be asymmetries in how the absolute basis behaves depending on whether the staking tokens trade at a premium or at a discount relative to the underlying asset, i.e., whether the liquid staking basis is positive or negative. To capture these, we include the binary variable *Liquid Staking Basis*$^+$ which is equal to 1 if the basis is positive.

Fourth, we control for the effect of the Merge. While this event itself did not immediately allow users to collect their staking rewards or un-stake their tokens, it was an assurance that the future upgrades (the “Capella” and “Shanghai” upgrades) would be successfully implemented. To control for the impact of the reduction in uncertainty regarding PoS adoption after the Merge and the subsequent increase attractiveness of liquid staking, the control vector includes the binary variable *Post Merge* indicating the period after the Merge of the Ethereum Mainnet with its PoS blockchain from September 15th, 2022 onward.

Hour captures fixed effects for each hour of the day to account for differences in reward payout mechanics. For example, in stETH, staking rewards are paid out once a day, which potentially leads to intraday price increases relative to the underlying asset due to accrued rewards.

Additionally, although liquid staking tokens serve a different economic purpose than cryptocurrency futures contracts, their pricing relative to the underlying asset might be similar in some respects. For example, limits to arbitrage and behavioral factors might similarly impact investors in the cryptocurrency futures market and in the liquid staking market. Furthermore, the liquid staking and cryptocurrency futures bases might co-move due to common risk premia as shown for the commodity futures market by Bailey and Chan (1993). We study this co-movement of the futures-spot and the liquid staking bases of the same underlying asset from the same venue by including the *Futures Basis* as given by the absolute value of the price difference between the perpetual futures contract and the underlying asset in the spot market, relative to the price in the spot market.

A noteworthy point is that, similar to the existing literature on price differences between derivative and spot markets (Roll, Schwartz, and Subrahmanyam 2007), there may be simultaneity, specifically between the basis, staking rewards, and liquidity. For example, large differences between spot and derivative prices may lead to arbitrage trading which in turn may increase liquidity, leading to an underestimation of the effect of liquidity on the basis. In a similar vein, liquidity may be high if relative staking rewards are large, attracting traders. Lacking reasonably exogenous shocks, in our main analyses we do not focus on these dynamics but instead concentrate on the simultaneous relationship of the variables of interest. However, in Appendix B we present supplementary results that document the dynamic relationship between the basis and limits to arbitrage.

To test for unit roots, we perform augmented Dickey-Fuller tests. The results reject the null hypothesis of a unit root for all variables used in the regressions. Furthermore, variance inflation factors indicate that multicollinearity is not a concern.

4.4 | Measuring Price Discovery

After having established the price determinants of liquid staking tokens, we investigate their contribution to the price discovery process of the underlying cryptocurrencies. We generally follow the approach of Alexander et al. (2020b) who compare price discovery between several Bitcoin futures and spot markets. In particular, we employ the information share measure by Lien and Shrestha (2009), which is a modified version of the measure by Hasbrouck (1995). One advantage of this modified measure is that it provides a unique estimate instead of upper and lower bounds of the share of price discovery. We supplement the analysis with the component share measure by Gonzalo and Granger (1995). Both measures are based on a decomposition of the permanent component and the cointegration errors of multiple cointegrated time series of prices using a multivariate vector error-correction model (VECM). For the lag length selection of the VECM, we rely on the Akaike information criterion, allowing for a maximum lag length of 40.

We use log prices sampled at the highest frequency of 15 seconds to compute the measures separately for each day and jointly for the full sample period, respectively. To compare the price discovery at the CEX and the DEX, we first need to make the implied stETH prices at the DEX comparable to the prices in USD at the CEX. We achieve this goal by converting the DEX implied prices to prices in USD using the stETH/ETH exchange rate from the CEX. While this step potentially introduces some noise as the DEX prices now partially reflect information from the CEX, the relative contribution of the staking derivatives to the underlying should only marginally be impacted by this.[15] However, to further address this concern, we additionally provide estimates for the price discovery of the liquid staking basis where no such adjustment is required. To this end, we apply the same price discovery measures to the liquid staking bases of stETH from both markets instead of applying them to price levels.

5 | Understanding the Liquid Staking Basis

In this section, we first present descriptive statistics, show the development of the liquid staking basis over time, and discuss the correlations between returns and trading volume of the various assets. We then investigate the determinants of the basis in a regression framework. For the descriptive statistics and the regressions, we first present the results for the DEX and then compare these results with those of the CEX. Finally, we compare the pricing of liquid staking tokens to the futures market.

5.1 | Descriptive Statistics and Correlations

5.1.1 | Descriptive Statistics for the DEX

Table 1 reports summary statistics of the key variables of our analysis. In Panel A, we present statistics on liquid staking via Lido Staked Ether. On average, there is a total value of USD 7.29 billion staked via this protocol, or 3.88 million Ether. The share of Lido in all Ethereum staking has increased throughout the sample, ranging from about 17% to almost one-third. This high concentration of staking in just one liquid staking provider raises concerns regarding limited decentralization. Average annualized staking rewards in Lido Staked Ether are 4.68%. For reference, when staking directly in Ethereum, annualized staking rewards are on average about 0.52 percentage points higher. In other words, staking via Lido costs about 10% of the staking rewards. There is also substantial time-series variation in staking rewards and in the reward difference, posing a risk to those staking their tokens.

Panel B shows summary statistics for the DEX, the Curve stETH/ETH liquidity pool. On average, there is USD 2.26 billion deposited in the pool. Since slightly less than half of this is in the form of ETH and the rest in stETH, a back-of-the-envelope calculation suggests that on average about 15.5% of all stETH is deposited in the pool. The average hourly trading volume is USD 1,230, although there is substantial time-series variation as can be seen in the maximum value and the kurtosis. The sum of deposits and withdrawals is close to zero on average, but likewise has some extreme values. For example, a withdrawal of USD 970 million occurred on May 12th, 2022, shortly after the collapse of TerraUSD, as investors were withdrawing founds across DeFi protocols and moving into safer assets.

Panel C reports that the absolute liquid staking basis at Curve is 96 basis points on average. The signed basis is almost always negative; the 95th percentile is negative while the maximum value of 0.08% is small in magnitude compared to the minimum

TABLE 1 | Descriptive statistics for Lido staking and DEX trading.

	Mean	SD	Min	P5	P50	P95	Max	Skew.	Kurt.
Panel A: Lido staking									
Total value locked	7.29	2.14	4.03	4.70	6.57	11.73	13.12	0.9	2.7
Staked amount	3.88	1.44	1.38	1.48	4.15	6.00	6.95	−0.4	2.2
Lido share of staking	27.64	5.26	16.98	17.49	30.03	32.16	32.62	−1.1	2.5
Staking rewards	4.68	0.92	3.57	3.83	4.62	6.04	10.21	2.8	15.5
Δ Staking rewards	0.52	0.10	0.40	0.43	0.51	0.67	1.13	2.8	15.5
Panel B: Curve stETH/ETH pool									
Pool size	2.26	1.58	0.56	0.73	1.44	5.22	6.12	0.9	2.3
Liquidity utilization	0.09	0.39	0.00	0.00	0.01	0.33	13.31	14.0	302.9
Pool fees	0.18	0.39	0.00	0.01	0.07	0.71	5.24	6.5	59.6
Trading volume	1.23	6.55	0.00	0.00	0.14	4.63	311.43	21.3	675.2
Net liquidity flow	−0.12	17.79	−970.13	−4.39	−0.00	4.35	392.98	−12.7	797.1
Panel C: Liquid staking basis									
Liquid staking basis$_{\text{signed}}$	−0.95	1.25	−6.85	−3.59	−0.41	−0.01	0.08	−1.8	6.1
Liquid staking basis	0.96	1.25	0.00	0.02	0.41	3.59	6.85	1.8	6.1

Note: This table shows summary statistics using data at 1 hour intervals. *Total Value Locked* is the value of all Ether staked via Lido Staked Ether in USD 1 billion. *Staked Amount* is the amount of all Ether staked via Lido Staked Ether in the 1 million Ether. *Lido Share of Staking* is the fraction of Ether staked via Lido to all Ether staked in the Ethereum network in percent. *Staking Rewards* is the annualized yield of staking Ether via Lido in percent and Δ *Staking Rewards* the difference between yields for staking Ether directly and staking via Lido in percentage points. *Pool Size* is the total value deposited in the pool in USD 1 billion. *Liquidity Utilization* is the ratio of trading volume in the pool to the total size of the pool in basis points. *Pool Fees* are the fees paid to liquidity providers, expressed as annualized yields in percent. *Trading Volume* is the value exchanged in the pool in USD 1k. *Net Liquidity Flow* is the net sum of the pool's in- and outflows in USD 1mn. *Liquid Staking Basis* is the absolute relative price difference between the staking token and the underlying cryptocurrency relative to the price of the underlying cryptocurrency in percent and *Liquid Staking Basis*$_{\textit{signed}}$ the corresponding signed value.

of −6.85%. This indicates that stETH typically trades at a discount relative to its underlying asset ETH.

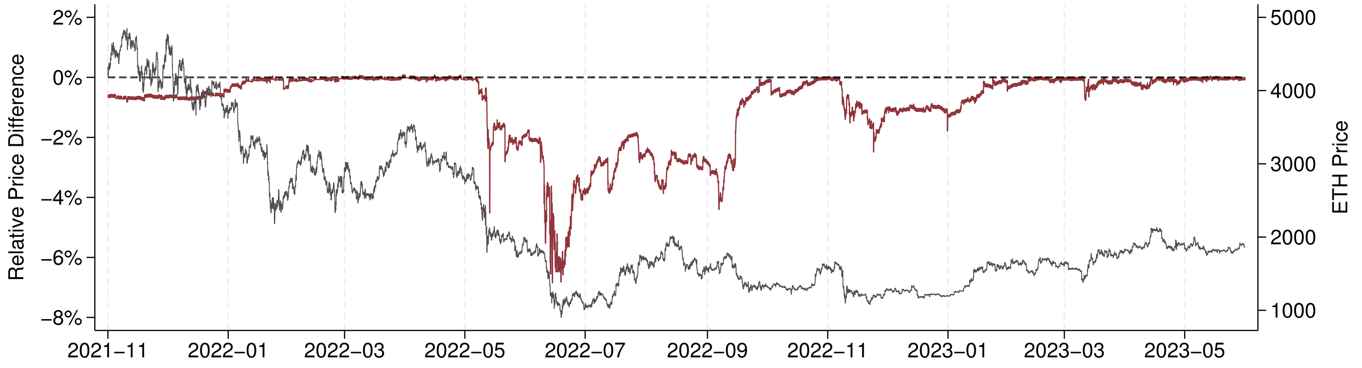
There is considerable time-variation in the liquid staking basis as shown in Figure 2. stETH began trading at a substantial discount relative to ETH during the crypto winter, starting around the collapse of TerraUSD in May 2022. The discount amid the general turmoil in cryptocurrency markets may reflect trading imbalances due to liquidity needs by failing cryptocurrency financial institutions such as Celsius Network, a cryptocurrency lending company, or Three Arrows Capital, a cryptocurrency hedge fund, that had to sell large quantities of stETH in the secondary market, in particular using the Curve liquidity pool (see, e.g., Nansen 2022). Since at that time, ETH could not be un-staked until the future Shanghai upgrade to the Ethereum network, the overall supply of stETH could not decrease. The only way for traders to exit their stETH positions was to sell them in the secondary market, on which counterparties appear to have required a discount as compensation for potential illiquidity or other risks. Since the successful Merge of the PoS chain to the Ethereum Mainnet and the subsequent reduction in uncertainty surrounding Ethereum's transition from PoW to PoS, the basis is reduced in magnitude. However, even after the Merge, stETH sometimes trades at a significant discount, for instance during adverse market conditions such as those following the collapse of FTX in November 2022. Overall, there is substantial time-variation and volatility clustering in the basis of stETH that is indicative of basis risk. We investigate potential determinants of this time-variation below.

Figure 2 also shows the development of the total value deposited in the Curve liquidity pool. Generally, liquidity in the pool has decreased during the sample period, decreasing from about USD 6 billion in November 2021 to about USD 1 billion in May 2023. The largest declines in pool size coincide with shocks to the wider cryptocurrency market, most prominently the depegging of TerraUSD in May 2022 but also the collapse of FTX in November 2022. These periods also show the highest levels of trading activity.

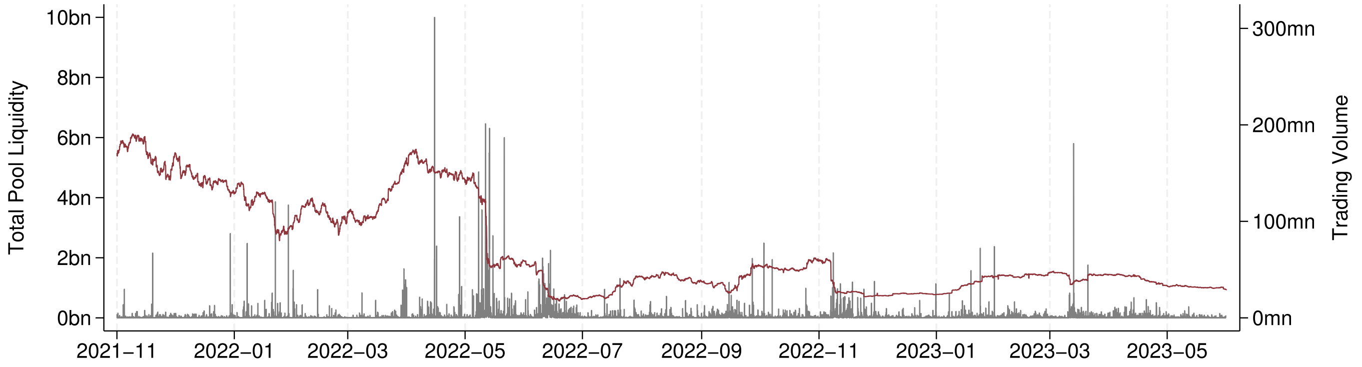
5.1.2 | Descriptive Statistics for the CEX

The CEX data allows us to compare trading activity in different assets traded at a single exchange. We present descriptive statistics in Table 2. Firstly, we observe that the staking tokens' returns have more downside risk than those of the underlying assets. This is shown by the minimum hourly return of −19.04% for stETH which is almost twice as high as that of its underlying asset ETH. These fatter tails are further reflected in higher kurtoses.

Secondly, trading volume is orders of magnitude larger in the underlying cryptocurrencies, reflecting the age and relative maturity of the spot market in addition to the broader acceptance of the underlying cryptocurrencies within DeFi applications compared to the liquid staking derivatives. Compared to trading at the DEX, trading activity in the liquid staking tokens at the CEX is lower. Similarly, secondary market liquidity is lower in the staking tokens, as evident on average bid-ask spreads that are between 4 and 5 times wider than that of the underlying assets. However, the average bid-ask spreads of about 4 basis points are still low compared to other financial markets such as large-cap equities (see e.g. Aspris et al. 2022).



(a) Liquid Staking Basis (stETH)



(b) DEX pool liquidity and trading volume

FIGURE 2 | Liquid staking basis and DEX pool dynamics. *Note:* These graphs show the hourly development of the stETH-ETH liquidity pool at the decentralized exchange Curve. Panel A shows the signed liquid staking basis based on transactions in the pool. The gray line shows the price of Ether in USD taken from the Exchange Kraken. Panel B shows the total liquidity available in the pool in USD and the hourly trading volume. [Color figure can be viewed at wileyonlinelibrary.com]

Thirdly, the liquid staking basis is higher on average for Solana-based tokens than for Ether-based tokens, although the opposite holds for the volatility of the basis.[16] For stETH, the liquid staking basis is about 40% larger at the CEX than at the DEX. Finally, the liquid staking basis is generally larger in magnitude and more volatile than the futures basis, suggesting that the market for liquid staking tokens is less mature and not as well integrated as the futures market.

We plot the development of the liquid staking basis as traded at the CEX in Figure 3. Overall, the graph for stETH closely resembles the one for the same asset at the DEX in Figure 2. A striking feature of the graphs for the Solana-based tokens is the linear increase in the basis. These liquid staking tokens are trading at a premium relative to the underlying cryptocurrency Solana because they are reward-bearing, i.e., the staking rewards mechanically increase their value, explaining the higher level of the basis in the descriptive statistics above. In the subsequent analyses, we hence use the de-trended basis. Since the payout mechanism differs for the rebasable stETH, there is no upward trend for this token. While stETH suffered from a substantial de-pegging starting May 2022 as discussed above, the Solana-based tokens do not show this pattern. A potential explanation for this finding is that in Solana, un-staking was already possible and the supply of the tokens could hence adjust.

5.1.3 | Correlations

We present pair-wise time series correlations in Figure 4. Panel (a) shows correlations of returns of the underlying assets, the staking tokens, and the futures contracts. The correlation between spot and futures returns is practically 100%, indicating that the spot and futures markets are highly integrated and efficient. The correlations of the staking tokens' returns with those of the underlying assets are high, but overall lower than for the futures market. For Ether the correlation is 93% while for Solana the correlations range from 95% to 98%. This suggests that, while the prices of liquid staking tokens generally closely follow the prices of the underlying assets, they do so less perfectly than futures contracts. This could be due to lower levels of high-frequency arbitrage trading activity in the market for staking tokens, potentially because of limits to arbitrage.

The results for the correlations in trading volume in panel (b) generally mimic those for returns, albeit at an overall lower level. The correlation is strongest for the trading volume in the spot and futures markets. For the staking tokens, the correlation of trading volume with that of the spot market ranges from 16% to 20%, indicating that there is a common component in trading activity between spot and liquid staking markets.

TABLE 2 | Descriptive statistics for trading at the CEX.

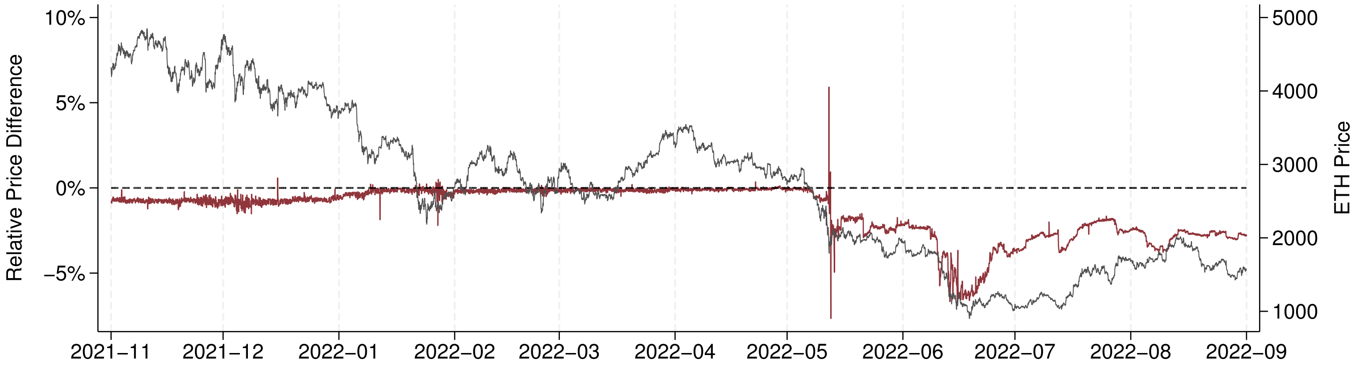
	Mean	SD	Min	P5	P50	P95	Max	Skew.	Kurt.
Panel A: Lido staked ether									
Return	−0.01	1.00	−19.04	−1.45	0.00	1.40	10.90	−0.9	30.1
Volume	0.01	0.07	0.00	0.00	0.00	0.05	1.97	15.2	315.9
Spread	4.16	5.06	0.00	0.99	2.83	10.40	94.32	6.6	71.3
Liquid staking basis$_{\text{signed}}$	−1.33	1.47	−7.66	−3.84	−0.71	−0.03	5.92	−1.2	4.1
Liquid staking basis	1.35	1.46	0.02	0.07	0.71	3.85	7.68	1.3	4.1
Panel B: Ether									
Return	−0.01	0.95	−10.40	−1.45	−0.01	1.34	7.37	−0.1	11.5
Volume	20.97	19.95	1.29	4.81	15.11	56.43	320.14	3.8	29.8
Spread	0.82	0.79	0.03	0.19	0.61	2.06	14.73	4.7	47.0
Futures basis$_{\text{signed}}$	0.00	0.02	−0.09	−0.03	0.00	0.04	0.21	0.8	6.2
Futures basis	0.02	0.01	0.01	0.01	0.02	0.04	0.21	3.0	20.2
Forced liquidations	0.38	0.51	0.00	0.04	0.21	1.23	5.57	4.8	39.4
Panel C: Marinade staked solana									
Return	−0.03	1.29	−21.43	−1.99	0.00	1.93	11.68	−0.4	17.8
Volume	0.01	0.05	0.00	0.00	0.00	0.07	1.65	12.8	295.6
Spread	3.76	4.03	0.00	0.72	2.70	10.61	97.32	5.0	60.7
Liquid staking basis$_{\text{signed}}$	3.79	1.43	0.54	1.47	3.80	5.95	6.75	−0.1	1.9
Liquid staking basis	3.79	1.43	0.71	1.47	3.80	5.95	6.75	−0.1	1.9
Panel D: Lido staked solana									
Return	−0.03	1.32	−23.97	−2.03	0.00	1.89	14.43	−0.4	25.4
Volume	0.01	0.08	0.00	0.00	0.00	0.05	3.80	32.0	1319.7
Spread	4.52	5.52	0.00	0.76	2.99	13.57	98.11	5.3	50.9
Liquid staking basis$_{\text{signed}}$	3.03	1.41	−5.18	0.91	2.87	5.21	8.45	0.1	2.1
Liquid staking basis	3.04	1.40	0.32	0.92	2.87	5.21	10.22	0.2	2.0
Panel E: Solana									
Return	−0.03	1.31	−21.15	−2.01	−0.03	1.96	12.94	−0.3	17.7
Volume	6.08	5.60	0.26	1.39	4.46	16.50	108.59	3.8	36.0
Spread	1.33	1.25	0.10	0.38	1.03	3.21	27.96	6.1	76.1
Futures basis$_{\text{signed}}$	−0.00	0.03	−0.41	−0.05	−0.00	0.04	0.14	−1.8	18.5
Futures basis	0.03	0.02	0.01	0.01	0.02	0.06	0.41	5.5	55.8
Forced liquidations	0.33	0.52	0.00	0.03	0.18	0.93	5.38	6.0	50.6

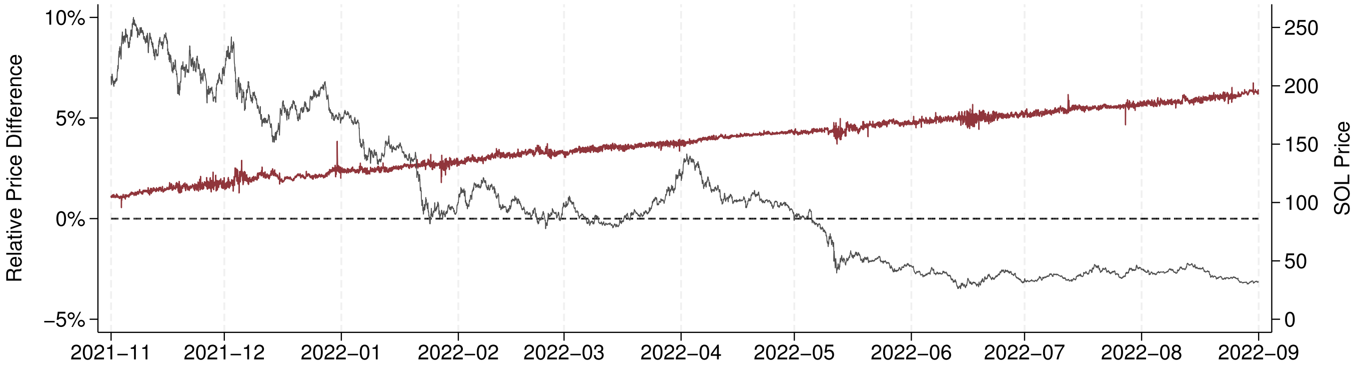
Note: This table shows summary statistics using data at 1 hour intervals for assets traded at the centralized exchange FTX. *Return* is the log return based on closing prices in percent. *Volume* is the trading volume in USD 1k. *Spread* is the estimated bid-ask spread in basis points. *Liquid Staking Basis* is the relative price difference between the staking token and the underlying cryptocurrency relative to the price of the underlying cryptocurrency in percent and *Liquid Staking Basis*$_{abs}$ the corresponding absolute value. The *Futures Basis* is computed analogously between the perpetual futures contracts and the underlying cryptocurrency. *Forced Liquidations* is the daily fraction of trading volume in all assets for a given underlying cryptocurrency (i.e., staking derivatives, futures, and spot market) resulting from automatic unwinding of traders' positions due to margin calls in percent.

Finally, we show the correlations in the (de-trended) absolute basis in panel (c). Importantly, the correlation between the liquid staking basis of stETH as traded on the DEX and the CEX is virtually perfect. This suggests that even though the CEX is substantially less liquid than the DEX for this particular asset, both bases strongly co-move. While we investigate this notion more formally in Section 6 when considering information shares, the high correlations documented here suggest that the basis is not trading venue specific. Instead, the basis and its determinants are likely closely related and the markets are integrated to a large extent, even though the level of the basis is significantly larger at the CEX. As expected, the correlation in the absolute basis across assets are much lower, ranging from practically 0% between stETH and stSOL to 34% between the two Solana-based tokens.

5.2 | Determinants of the Liquid Staking Basis

We now turn to the analysis of the potential determinants of the liquid staking basis. We first present the results for the DEX and then compare them to the CEX.


(a) Liquid Staking Basis (stETH)


(b) Liquid Staking Basis (mSOL)

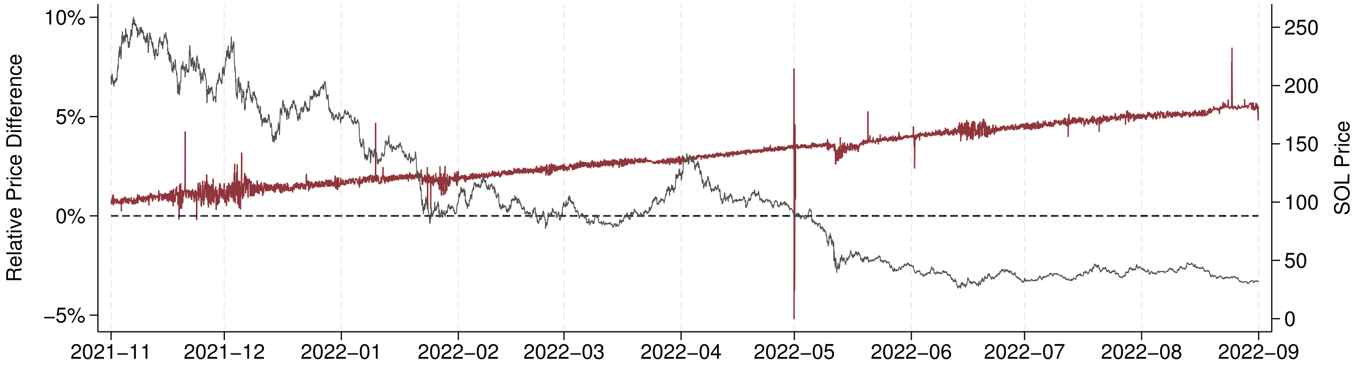

(c) Liquid Staking Basis (stSOL)

FIGURE 3 | Liquid staking basis at the CEX. *Note:* These graphs show the hourly development of the signed liquid staking basis at the centralized exchange FTX as given by the difference between the staking tokens and the underlying assets, relative to the price of the underlying. The thin gray line shows the price of the underlying cryptocurrency in USD. [Color figure can be viewed at wileyonlinelibrary.com]

5.2.1 | Determinants of the Liquid Staking Basis at the DEX

The regression results of Equation (5) for the DEX are reported in Table 3. In model (1), we only include the control variables for the crypto winter, the post Merge period, and the weekend. The liquid staking basis widens considerably by 2.7 percentage points during the crypto winter. After the Merge, the liquid staking basis is again close to its level before the TerraUSD collapse, as the sum of the two coefficients is close to zero. The successful Merge hence appears to have reduced uncertainty surrounding Ethereum's adoption of PoS that was reflected in the basis. We also find that the liquid staking basis is larger during the weekend. This result is consistent with the notion that institutional investors are better at exploiting potential arbitrage opportunities and are relatively more active during the week than on the weekend. The wider basis during the weekend may also indicate fluctuations in risk or risk preferences (Singal and Tayal 2020).

We then test Hypothesis 1 and add the difference in staking rewards between direct staking and staking via Lido in model (2). Higher values for Δ *Staking Rewards* indicate higher opportunity costs of staking with Lido compared to direct

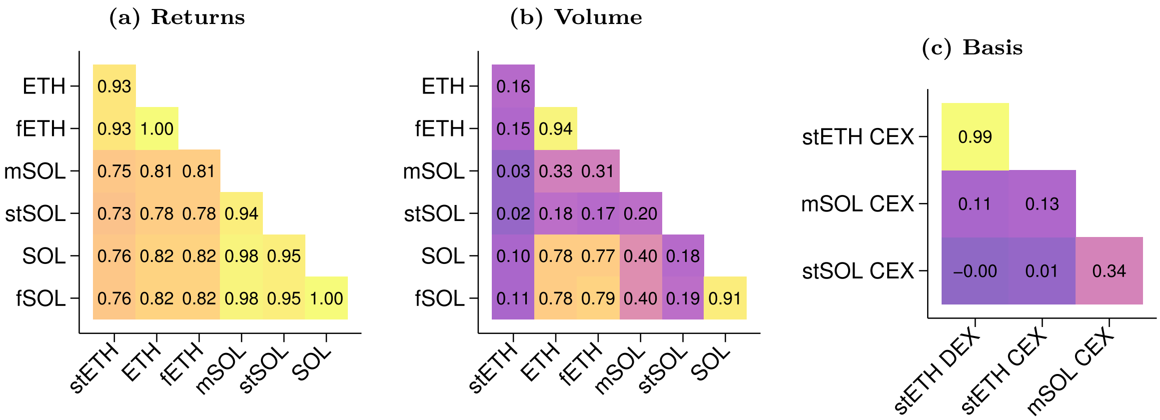


FIGURE 4 | Time-series correlations. *Note:* These graphs show time-series correlation coefficients. Figures (a) and (b) show correlations between hourly returns and trading volume in USD, respectively, for the liquid staking tokens (stETH, mSOL, and stSOL), the underlying cryptocurrencies (ETH and SOL), and the respective futures contracts (fETH and fSOL) as traded at the CEX. Figure (c) shows the correlations of the bases at the DEX and the CEX. [Color figure can be viewed at wileyonlinelibrary.com]

staking. Before the Merge, the effect of yield differences on the magnitude of the basis is positive and significant, suggesting that when Lido offers relatively low yields, the price discount for stETH is higher. While this finding supports Hypothesis 1, after the Merge the effect mostly disappears.[17]

The relative importance of Lido in Ethereum staking has a negative coefficient while not controlling for secondary market activity. However, when including variables relating to secondary market liquidity and trading activity in model (4), this effect turns positive. This finding is supportive of Hypothesis 2 and agrees with the voiced concerns by market participants regarding concentration risks. A high concentration in a single protocol potentially makes the entire network more vulnerable to attacks (Grandjean, Heimbach, and Wattenhofer 2023). Our findings suggests that traders may be aware of these risks and consider it in their trading decisions.

To study the volatility aspect of limits to arbitrage, model (3) includes the proxy for volatility of the underlying asset and the corresponding sign of returns. Consistent with Hypothesis 3, we find that higher volatility is associated with a wider liquid staking basis. However, the effect disappears when including the measures of liquidity in model (4), likely because increases in volatility are oftentimes associated with decreases in liquidity. Note that all models include the control variable for the crypto winter, which might partially capture the associated increase in overall market volatility. The direction of price movements does not appear to influence the basis.

To further investigate the liquidity aspect of limits to arbitrage, we include variables related to liquidity and trading activity in model (4). As expected, the liquidity available in the Curve pool is negatively related to the magnitude of the basis. A one percent increase in pool size is associated with a decrease in the basis of between 0.953 and 1.205 basis points. Trading volume is not significantly associated with the liquid staking basis, nor is the contemporaneous net flow of liquidity to and from the Curve pool. However, the coefficient for trading fees in the Curve pool is positive and highly significant. These results suggest that times of low liquidity and high transaction costs are associated with a wider basis. This finding supports Hypothesis 3 and is consistent with the notion that limits to arbitrage in the secondary market hinder the exploitation of any mispricing, similar to the futures-cash basis in Roll, Schwartz, and Subrahmanyam (2007).[18]

Model (5) then includes behavioral factors. An increase in investor attention is associated with a highly significant decrease in the basis. The significantly negative coefficient of investor attention suggests that when investors pay more attention towards liquid staking, price differences tend to be smaller in magnitude. The difference in the basis between days with virtually no and days with the highest Google search volume is about 25 basis points. The results thus confirm for liquid staking tokens what Han, Li, and Yin (2017) find for commodity futures contracts: Higher investor attention corresponds to fewer arbitrage opportunities as more investors monitor the markets and take advantage of any arising prices differences. Investor sentiment is likewise an important determinant of the liquid staking basis. An increase in the “fear index”, i.e., an increase in the implied volatility of cryptocurrency derivatives, is associated with a wider basis. This is consistent with lower risk appetite and decreased risk absorption capacities of potential arbitrageurs during times of low sentiment (Cheng, Kirilenko, and Xiong 2015). Overall, these results support Hypothesis 4 regarding the impact of investors' attention and sentiment on the liquid staking basis.

In model (6), we additionally control for the sign of liquid staking basis. The coefficient is negative and highly statistically significant. This indicates that there is an asymmetry in the basis in the sense that the basis is smaller in magnitude when it is positive than when it is negative.

Finally, we repeat the analysis of model (6) in model (7) while ending the sample in August 2022, before the collapse of FTX and thus also before the Merge. Overall, we find that the results are similar. The exception is the insignificant coefficient for the crypto winter, an effect likely subsumed by the variable for investor sentiment. We present these results to not only show the robustness of our findings, but also because this sample

TABLE 3 | Determinants of the liquid staking basis at the DEX.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Crypto winter	2.683***	3.127***	3.107***	1.257***	1.181***	1.153***	0.033
	(131.71)	(113.92)	(113.17)	(31.00)	(26.92)	(26.29)	(0.63)
Post merge	−2.510***	−1.555***	−1.505***	0.290	−0.910***	−0.928***	
	(−120.69)	(−5.33)	(−5.11)	(1.36)	(−4.20)	(−4.28)	
Weekend	0.021*	0.027**	0.036***	0.026***	0.027***	0.027***	0.055***
	(1.82)	(2.43)	(3.17)	(2.90)	(3.18)	(3.17)	(4.25)
Δ Staking rewards		2.539***	2.551***	5.922***	2.745***	2.680***	23.596***
		(3.89)	(3.88)	(12.54)	(5.73)	(5.59)	(37.85)
Δ Staking rewards × post merge		−2.348***	−2.396***	−6.180***	−3.004***	−2.937***	
		(−3.55)	(−3.61)	(−12.92)	(−6.21)	(−6.07)	
Lido share of staking		−0.027***	−0.027***	0.015***	0.022***	0.024***	0.267***
		(−4.39)	(−4.22)	(3.10)	(4.40)	(4.69)	(36.45)
$\text{Return}^{\text{Absolute}}_{\text{Underlying}}$			6.391***	0.826	−1.673*	−1.713*	−2.979**
			(3.97)	(0.76)	(−1.73)	(−1.77)	(−2.52)
$\text{Return}^{\text{Positive}}_{\text{Underlying}}$			−0.009	−0.003	−0.004	−0.004	−0.012
			(−0.91)	(−0.38)	(−0.47)	(−0.46)	(−1.01)
Pool size				−1.205***	−1.097***	−1.097***	−0.953***
				(−54.33)	(−59.11)	(−59.23)	(−40.35)
Trading volume				−0.001	−0.001	−0.001	−0.001
				(−1.06)	(−1.19)	(−0.91)	(−0.43)
Net liquidity flow				0.088	0.121	0.169	0.109
				(0.26)	(0.31)	(0.45)	(0.23)
Pool fees				0.391***	0.354***	0.350***	0.307***
				(14.57)	(14.03)	(14.04)	(10.95)
Attention					−0.252***	−0.251***	−0.073**
					(−15.18)	(−15.22)	(−2.14)
Sentiment					2.917***	2.957***	4.905***
					(12.42)	(12.62)	(23.32)
Liquid staking basis_+						−0.291***	−0.183***
						(−13.23)	(−5.50)
Hour FE	✓	✓	✓	✓	✓	✓	✓
Adj. R^2	0.75	0.76	0.76	0.86	0.86	0.87	0.89
Observations	13,837	13,789	13,788	13,773	13,773	13,773	7,280
Sample end	2023-05	2023-05	2023-05	2023-05	2023-05	2023-05	2022-08

Note: This table shows time-series regression results for the determinants of the absolute percentage liquid staking basis as implied by the exchange rate between the Lido Staked Ether and the underlying cryptocurrency Ether as traded at the decentralized exchange Curve in percent using hourly data. *Crypto Winter* is an indicator variable for the period starting May 9th, 2022. *Post Merge* is an indicator variable for the period starting September 15th, 2022. *Weekend* is an indicator variable for Saturdays and Sundays (UTC). Δ *Staking Rewards* the difference between yields for staking Ether directly and staking via Lido in percentage points. *Lido Share of Staking* is the fraction of Ether staked via Lido to all Ether staked in the Ethereum network in percent. $Return^{Absolute}_{Underlying}$ is the absolute log return of the underlying in percent and $Return^{Positive}_{Underlying}$ an indicator variable set to one if this return is positive. *Pool Size* is the logarithm of the total USD value deposited in the pool. *Trading Volume* is the value exchanged in the pool in USD 1k. *Net Liquidity Flow* is the net sum of the pool's in- and outflows in USD 1mn. *Pool Fees* are the fees paid to liquidity providers, expressed as annualized yields in percent. *Attention* is the daily Google search volume index for the term "liquid staking", scaled to the interval of 0 to 1. *Sentiment* is the value of the Deribit Volatility (DVOL) Index which tracks the implied volatility of cryptocurrency derivatives. *Liquid Staking* $Basis_+$ is an indicator variable set to 1 if the signed liquid staking basis is positive. Robust t-statistics are given in parentheses. ***, **, * denotes significance at the 1%, 5%, 10%–level, respectively.

TABLE 4 | Determinants of the liquid staking basis at the CEX.

	(1)	(2)	(3)	(4)	(5)	(6)
Crypto winter	2.586***	2.002***	1.941***	1.934***	0.888***	0.888***
	(116.64)	(49.05)	(47.56)	(47.85)	(17.82)	(18.04)
Weekend	0.052***	0.043**	0.060***	0.070***	0.096***	0.095***
	(2.88)	(2.52)	(3.52)	(4.16)	(7.33)	(7.22)
Δ Staking rewards		23.795***	24.216***	23.654***	24.678***	24.586***
		(23.40)	(23.69)	(23.33)	(28.95)	(29.15)
Lido share of staking		0.193***	0.199***	0.199***	0.310***	0.309***
		(19.36)	(19.87)	(20.08)	(32.99)	(33.31)
$\text{Return}^{\text{Absolute}}_{\text{Underlying}}$			12.841***	10.750***	0.455	0.646
			(6.20)	(5.34)	(0.31)	(0.45)
$\text{Return}^{\text{Positive}}_{\text{Underlying}}$			0.026*	0.009	0.009	0.009
			(1.65)	(0.56)	(0.66)	(0.69)
$\text{Spread}_{\text{Staking token}}$				1.637***	0.776***	0.777***
				(9.40)	(4.83)	(4.79)
$\text{Volume}_{\text{Staking token}}$				1.115***	1.008***	1.011***
				(3.76)	(4.50)	(4.49)
Forced liquidations				−0.022	0.057***	0.057***
				(−1.53)	(4.18)	(4.20)
Attention					−0.086**	−0.085**
					(−2.21)	(−2.19)
Sentiment					7.661***	7.688***
					(35.08)	(35.21)
$\text{Liquid staking basis}_{+}$						−0.078***
						(−2.94)
Futures basis						−0.783
						(−0.98)
Hour FE	✓	✓	✓	✓	✓	✓
Adj. R^2	0.74	0.78	0.78	0.79	0.85	0.85
Observations	7,296	7,296	7,295	7,295	7,295	7,295
Sample end	2022-08	2022-08	2022-08	2022-08	2022-08	2022-08

Note: This table shows time-series regression results for the determinants of the absolute liquid staking basis for Lido Staked Ether and the underlying cryptocurrency Ether as traded at the centralized exchange FTX using hourly data. The independent variables are defined as in Table 2 and Table 3 except *Spread* which is here given in percentage points. All models include a constant and fixed effects for the hour of the day. Robust t-statistics are given in parentheses. ***, **, * denotes significance at the 1%, 5%, 10%–level, respectively.

period aligns with the sample used in the analyses of the CEX. The DEX and CEX results can hence be more directly compared.

5.2.2 | Determinants of the Liquid Staking Basis at the CEX

We now analyze the potential determinants of the liquid staking basis at the CEX in Table 4. Overall, we confirm that the liquid staking bases at the CEX and at the DEX are influenced by the same economic factors. In particular, the basis is substantially wider during the crypto winter, during the weekend, when opportunity costs of staking via Lido are relatively high, and when Ethereum staking is more strongly concentrated in the Lido protocol. These findings are consistent with those for the DEX and similarly support Hypothesis 1 and Hypothesis 2.

Volatility as measured by absolute returns is positively associated with the magnitude of the basis. However, its impact becomes insignificant as we include variables that capture aspects of risk, particularly the implied volatility index that is used to measure sentiment. Secondary market liquidity, although measured differently for the CEX, is inversely related to the basis, similar to that of the DEX. Specifically, estimated bid-ask spreads are highly significantly and positively associated with the basis. These results further confirm Hypothesis 3 and imply that limits to arbitrage are an important factor for explaining the basis. Perhaps unexpectedly, the magnitude of the liquid staking basis also increases in the trading volume of the staking token. Since we already control for liquidity by including bid-ask spreads, this finding is consistent with the notion that high volume is indicative of noise trader risk, making arbitrage activity more difficult as in De Long et al. (1990).[19]

We further investigate the impact of funding liquidity on the liquid staking basis. Periods when forced liquidations are relatively frequent coincide with periods when the basis is relatively wide while controlling for market sentiment. Forced liquidations indicate lower funding liquidity and more capital-constrained investors. This finding is thus consistent with the notion that when funding liquidity is low, some capital-constrained investors quickly need to liquidate their positions in the staking derivatives, hence decreasing the tokens' prices relative to the underlying asset. Likewise, funding constraints also make arbitrage activity more difficult and expensive, further hindering the exploitation of arbitrage opportunities.[20]

Regarding investor behavior, similar to the DEX and in support of Hypothesis 4, we find evidence that increased awareness of liquid staking is associated with smaller price differences while investor “fear” is positively correlated with the magnitude of the basis. The asymmetry in the magnitude of the liquid staking basis depending on its sign is likewise present at the CEX.

Finally, we relate the liquid staking basis to the futures-spot basis. The coefficient for the futures basis is statistically insignificant. Therefore, the results do not indicate that the futures basis conveys any additional information for the magnitude of the liquid staking basis.

5.3 | Differences to the Determinants of the Futures-Spot Basis

To gain a deeper understanding of the liquid staking basis, in this section we examine how it compares to that of perpetual cryptocurrency futures contracts. To this end, we estimate the same model as before but using the futures basis as the dependent variable. We then compare the results to those of the liquid staking basis. This approach allows us to test whether the determinants of price differences are similar across markets. The results can be found in Table 5. For brevity, we only highlight several key findings.

Expectedly, both limits to arbitrage and investor behavior influence the basis of the perpetual futures contracts, consistent with the prior literature on the pricing of futures contracts. While we only find a weakly significant effect of volatility on the absolute basis, the coefficient for bid-ask spreads is positive and highly significant. Investor fear is likewise positively associated with the width of the basis, although the magnitude of this effect is much smaller than for liquid staking tokens. Furthermore, the futures-spot basis is also wider during the weekend when controlling for liquidity.

However, there are also important differences regarding the determinants of both bases. In particular, investor attention is positively associated with the futures-spot basis, suggesting that in the more mature futures market, a lack of investor attention is generally not a contributing factor to mispricing. On the contrary, in that market, investor attention may be related to investor herding or the formation of bubbles, which make arbitrage activity more difficult (see also Cretarola and Figà-Talamanca 2020). Moreover, the coefficient for forced liquidations is negative and significant in this market. A potential partial explanation for this result could be that, contrary to the staking tokens, the futures sometimes trade at a premium while the staking tokens usually trade at a discount. A decrease in funding liquidity might thus still depress futures prices relative to the underlying asset, but when futures generally trade at a premium, the absolute basis would then decrease.

To summarize, the determinants of the bases of liquid staking tokens and of futures contracts exhibit several similarities. Still, there are some differences, likely due to the relative maturity of the futures market and the different economic purposes the tokens serve.

6 | Price Discovery in Liquid Staking Tokens

Without liquid staking, the staked amount is taken out of circulation and hence cannot be used for trading. This illiquidity may thus impair informational efficiency and price discovery. Conversely, the liquidity provided by liquid staking tokens may facilitate price discovery, similar to how the liquidity provided by cryptocurrency futures benefits the price discovery process (Aleti and Mizrach 2021; Alexander et al. 2020b). We formally investigate this notion using modified information shares (Lien and Shrestha 2009) and component shares (Gonzalo and Granger 1995).

The results for Lido Staked Ether including both the DEX and the CEX can be found in Panel A of Table 6. Overall, the staking tokens contribute a substantial amount to price discovery. The median of the daily modified information share of the staking tokens at the DEX is about 23%, while the mean and the estimated share for the full sample are even higher. Likely due to the relatively low liquidity at the CEX, the staking tokens traded there only contribute little to the price discovery process. Component shares typically confirm these findings, although the share of stETH at the DEX for the overall sample is even larger.

In Panel B, we only compare the liquid staking basis at the CEX and the DEX. Both information shares and component shares are much larger for the liquid staking basis at the DEX, suggesting that it leads price discovery, even though we have previously documented that the correlation of the bases is at almost 100%. This result aligns with the overall higher liquidity and trading activity at the DEX as compared to liquid staking tokens as traded on the CEX.

The previous results mask the substantial time-variation in the staking tokens' contributions to the price discovery process. To shed more light on this aspect, we plot the evolution of the price

TABLE 5 | Determinants of the futures basis.

	(1)	(2)	(3)	(4)
fETH × crypto winter	0.001**	−0.003***	−0.003***	−0.002***
	(2.24)	(−7.40)	(−6.65)	(−6.44)
fSOL × crypto winter	0.006***	0.003***	0.003***	0.004***
	(9.33)	(5.57)	(6.36)	(7.76)
Weekend	−0.002***	0.001**	0.001**	0.001**
	(−6.04)	(2.12)	(2.09)	(2.02)
$\text{Return}^{\text{Absolute}}_{\text{Underlying}}$		0.001**	0.001*	0.001*
		(1.98)	(1.88)	(1.90)
$\text{Return}^{\text{Positive}}_{\text{Underlying}}$		0.000	0.000	0.000
		(0.35)	(0.25)	(0.21)
$\text{Spread}_{\text{Futures}}$		0.877***	0.831***	0.837***
		(13.57)	(11.91)	(12.00)
$\text{Volume}_{\text{Futures}}$		−0.019***	−0.018***	−0.019***
		(−3.61)	(−3.36)	(−3.55)
Forced liquidations		−0.001***	−0.001***	−0.001***
		(−4.46)	(−3.47)	(−3.55)
Attention			0.000***	0.000***
			(2.85)	(2.87)
Sentiment			0.018***	0.016***
			(5.90)	(5.45)
Futures basis$_+$				0.002***
				(7.34)
Contract FE	✓	✓	✓	✓
Hour FE	✓	✓	✓	✓
Adj. R^2	0.06	0.27	0.28	0.28
Observations	14,592	14,590	14,590	14,590

Note: This table shows regression results for the determinants of the absolute futures basis as given by the relative differences between the futures and the spot market in percent using hourly data from the CEX. *fETH* and *fSOL* are indicator variables for the respective futures contracts to capture the fixed effects. *Crypto Winter* is an indicator variable for the period starting May 9th, 2022. *Weekend* is an indicator variable for Saturdays and Sundays (UTC). $Return^{Absolute}_{Underlying}$ is the absolute log return of the underlying cryptocurrency in percent and $Return^{Positive}_{Underlying}$ an indicator variable set to one if this return is positive. *Spread* is the estimated bid-ask spread of the futures in percent. *Volume* is the trading volume in the futures contract in USD 1bn. *Forced Liquidations* is the daily fraction of trading volume in all assets for a given underlying cryptocurrency (i.e., staking derivatives, futures, and spot market) resulting from automatic unwinding of traders' positions due to margin calls in percent. *Attention* is the daily Google search volume index for the term “eth futures” or “sol futures”, scaled to the interval of 0 to 100. *Sentiment* is the value of the Deribit Volatility Index which tracks the implied volatility of cryptocurrency derivatives. All models include a constant and fixed effects for the hour of the day. Robust t-statistics are given in parentheses. ***, **, * denotes significance at the 1%, 5%, 10%–level, respectively.

discovery measures in Figure 5. As in Alexander et al. (2020b), the daily observations are exponentially smoothed to reduce noise. We again see that a significant amount of price discovery happens in the liquid staking tokens at the DEX. The relevance of this market for informational efficiency seems to have increased during the second half of the sample which covers the crypto winter. This suggests that the derivatives have become more important in the price discovery process over time, especially during the volatile period towards the end of the sample. Likewise, the contribution of the DEX to price discovery in the liquid staking basis has increased over time.

The results for the Solana-based tokens can be found in the Appendix. While we do not investigate DEX trading for these tokens, the conclusion that the importance of liquid staking tokens for the price discovery process has increased over time is confirmed in this market.

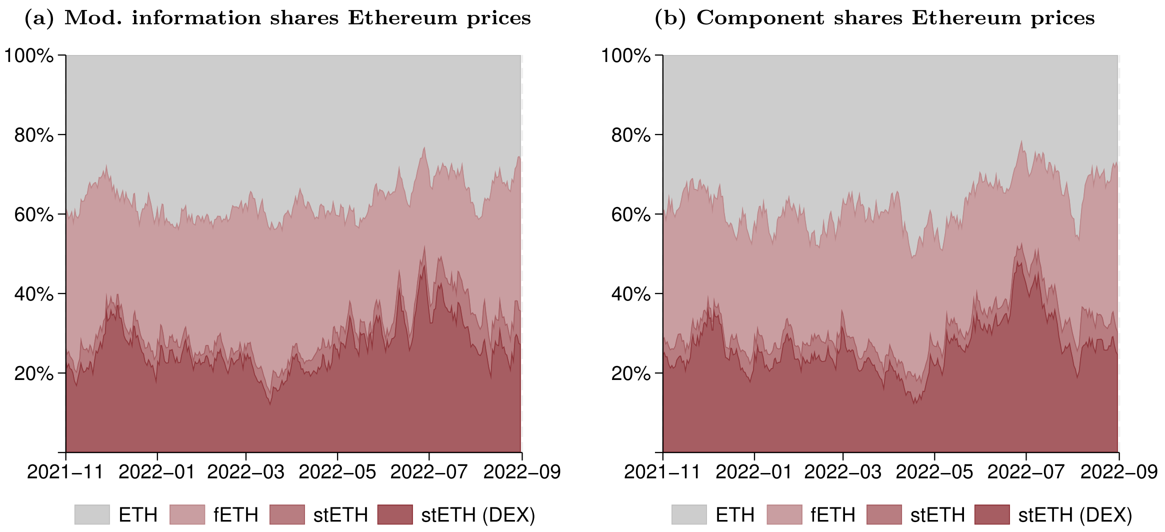
7 | Concluding Remarks

This paper examines a new form of derivative cryptocurrency token: liquid staking tokens, which represent a share of staked tokens in Proof-of-Stake blockchains. We empirically examine the liquid staking basis as given by the price differences between these tokens and their respective underlying cryptocurrency. While this basis is overall volatile, we find evidence that staking incentives, concentration risks, limits to arbitrage, and investor behavior constitute important factors that influence its magnitude. The basis is wider when the yields offered

TABLE 6 | Price discovery.

	Mod. Information Shares			Component Shares		
	Mean	**Median**	**Overall**	**Mean**	**Median**	**Overall**
Panel A: Ethereum prices						
stETH$_{\text{DEX}}$	26.93	22.93	53.90	26.83	24.04	87.41
stETH$_{\text{CEX}}$	4.53	2.57	1.48	4.60	3.50	0.82
fETH$_{\text{CEX}}$	32.70	31.23	20.99	31.80	28.91	2.67
ETH$_{\text{CEX}}$	35.85	37.15	23.63	36.77	35.68	9.11
Panel B: Ethereum liquid staking basis						
stETH basis$_{\text{DEX}}$	84.92	96.67	99.26	95.55	98.30	99.33
stETH basis$_{\text{CEX}}$	15.08	3.33	0.74	4.45	1.70	0.67

Note: This table shows modified information shares and component shares. Panel A is based on prices for the liquid staking token (stETH), the perpetual futures contracts (fETH), and the underlying spot market (ETH). For stETH, prices from both the CEX and the DEX is included. The other assets are traded at the CEX. Panel B is based on the liquid staking basis for stETH as traded on both venues. Means and medians are based on daily estimates while the column *Overall* contains the estimate for the full sample. All values are given in percent.


(a) Mod. information shares Ethereum prices **(b) Component shares Ethereum prices**

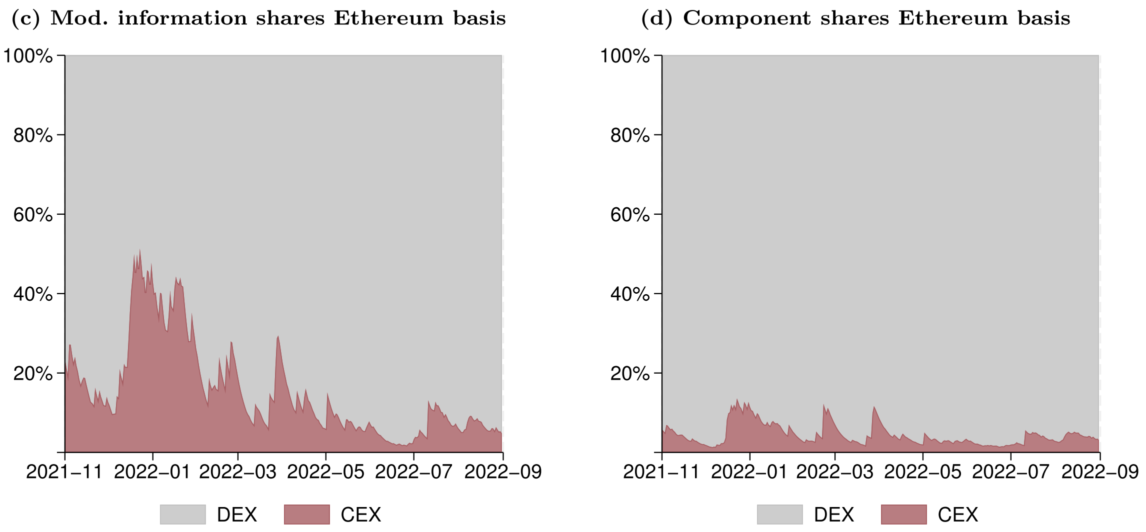

(c) Mod. information shares Ethereum basis **(d) Component shares Ethereum basis**

FIGURE 5 | Evolution of price discovery. *Note:* These graphs show daily modified information shares and component shares. For figures (a) and (b), the respective measures are shown for the prices of the liquid staking token (stETH), the perpetual futures contract (fETH), and the respective underlying asset (ETH). For stETH, prices from both the CEX and the DEX is included. The other assets are traded at the CEX. Figures (c) and (d) show the respective measures applied to the liquid taking bases from both trading venues. For readability, the information shares are smoothed with a trailing exponentially weighted moving average using 0.1 as the smoothing parameter. [Color figure can be viewed at wileyonlinelibrary.com] by the liquid staking protocol are low relative to the alternative of staking directly and when secondary market liquidity is low, consistent with limits to arbitrage. Concentration risks also appear to be a relevant factor for investors in this market, since the basis is wider when the share of stake deposited in the liquid staking protocol is larger. Conversely, the basis is narrower when investors pay more attention to liquid staking and when they are more optimistic regarding the future development of the cryptocurrency market. Information shares suggest that liquid staking tokens contribute a significant and growing amount to price discovery.

Our analyses are mostly based on contemporaneous relations. Future work may investigate the time dynamics of the determinants of the liquid staking basis in more detail or the predictability of future returns based on the basis. Similarly, most of our analyses focus on one liquid staking protocol, Lido Staked Ether. Even though at the time of writing, this protocol is by far the largest by market capitalization, and although we additionally consider two other liquid staking derivatives for robustness, future research may examine a broader set of staking tokens or compare liquid staking to other forms of delegated staking.

Our results have important implications for traders, other DeFi market participants, and financial regulators alike. An in-depth understanding of the pricing of these derivative tokens is essential for managing their risks. From a regulatory perspective, our results are relevant since liquid staking tokens appear to play a prominent role in the market turmoil starting in May 2022. Furthermore, the growing popularity of Proof-of-Stake blockchains and the switch of the currently second largest cryptocurrency network, Ethereum, from Proof-of-Work to Proof-of-Stake further amplify the relevance of our results.

Acknowledgments

We thank an anonymous referee, Daniele Bianchi, Roy Chen-Zhang, Andrew Urquhart, and conference and seminar participants at the 5th UWA Blockchain and Cryptocurrency Conference, the Cryptocurrency Research Conference 2022, the Cardiff Fintech Conference 2022, and the University of Mannheim for helpful comments and suggestions. All remaining errors are our own. Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that supports the findings of this study is available from the corresponding author upon reasonable request. Various sources were used to collect data for the analyses. Each source of data and materials has been pointed out throughout the paper. The FTX data is no longer publicly available, however, it is also available from the corresponding author upon reasonable request.

Endnotes

[1] More recently, options on cryptocurrencies have emerged and subsequently studied, see e.g. Hou et al. (2020), Cao and Celik (2021), and Alexander et al. (2023).

[2] A variation of PoS is delegated proof-of-stake (dPoS). Under this consensus mechanism, stakeholders can vote and elect witnesses from a limited list of potential witnesses (usually about 100) that then validate the next block of transactions. The weight of a vote increases in the amount of stake a participant has in the network. Rewards earned by the delegate are then shared with those that voted for this delegate. Delegates are partially incentivized by their reputation, as misbehavior makes it less likely to be voted again since the stakeholders potentially lose their stake.

[3] Direct staking is sometimes also referred to as “solo” or “protocol” staking.

[4] Examples of centralized exchanges offering such services include Coinbase and Binance.

[5] For a detailed taxonomy of liquid staking protocols, see Gogol et al. (2024).

[6] See https://lido.fi/ and https://marinade.finance/. Other popular liquid staking services are Ankr, Rocket Pool, and StakeWise.

[7] See https://ethereum.org/en/upgrades/

[8] https://coinmarketcap.com/currencies/steth/

[9] According to https://coinmarketcap.com/rankings/exchanges, FTX was the 3rd largest cryptocurrency exchange during our sample period. Importantly, it was the largest CEX for liquid staking tokens.

[10] Intervals of 15 seconds are the highest frequency of OHLCV data available via the FTX API.

[11] As discussed above, wstETH is a wrapped, reward-bearing version of stETH that reflects the staking rewards similarly to the Solana-based tokens in our sample. Lido stETH and wstETH tokens can be converted using a trustless smart contract. However, only stETH is listed at FTX, motivating our choice of asset.

[12] After estimating the model for stETH using the DEX data, we repeat the analysis using the CEX data for the same asset and compare the results to those obtained for the DEX. In analyses presented in the Appendix, we additionally include the two Solana-based liquid staking tokens in the model for robustness. Moreover, we estimate the model separately for each token. The overall conclusions do not materially differ.

[13] We use high and low prices within adjacent 15 second intervals and follow Corwin and Schultz (2012) by setting negative spread estimates to zero.

[14] All the times in this paper are UTC time.

[15] In untabulated tests, we repeat the analysis using 1 second quote midpoints from Coinbase, a very liquid and active exchange. We obtain virtually identical results.

[16] A noteworthy point here is that these statistics are based on the raw liquid staking basis, i.e., before any de-trending. For the reward-bearing Solana-based tokens, the basis here thus also reflects the staking rewards.

[17] Untabulated analyses show that the staking reward difference has changed after the successful Merge. Overall, the difference has increased and has become more volatile. This is because after the Merge, validators not only receive the consensus layer rewards, but also the more volatile execution layer rewards.

[18] In addition to the contemporaneous relationships studied in this section and to further examine the dynamics between limits to arbitrage and the basis, we conduct vector autoregressions and Granger causality tests similar to Roll, Schwartz, and Subrahmanyam (2007), Kadapakkam and Kumar (2013), and Han and Pan (2017). The results reported in Appendix B confirm our findings. As in these prior studies, there appears to be bidirectional Granger causality between liquidity and the magnitude of the basis.

[19] In untabulated analyses, we separate volume into trading volume when prices are increasing and when prices are decreasing. While the effect of volume is more positive when prices are decreasing, it is statistically significantly positive in both cases. Additionally, untabulated results show that additionally including trading volume of the spot market does not meaningfully change the result.

[20] Including forced liquidations, which are measured using data from the CEX, into the regressions for the DEX above yields almost identical results. When included in model (7) of Table 3, the coefficient for forced liquidations is 0.057 with a t-statistic of 4.03.

[21] Note that we only measure forced liquidations at a daily frequency, contrary to the other variables of the VAR analysis.

References

Aleti, S., and B. Mizrach. 2021. “Bitcoin Spot and Futures Market Microstructure.” *Journal of Futures Markets* 41: 194–225. https://doi.org/10.1002/fut.22163.

Alexander, C., J. Choi, H. R. Massie, and S. Sohn. 2020a. “Price Discovery and Microstructure in Ether Spot and Derivative Markets.” *International Review of Financial Analysis* 71: 101506. https://doi.org/10.1016/j.irfa.2020.101506.

Alexander, C., J. Choi, H. Park, and S. Sohn. 2020b. “Bitmex Bitcoin Derivatives: Price Discovery, Informational Efficiency, and Hedging Effectiveness.” *Journal of Futures Markets* 40: 23–43. https://doi.org/10.1002/fut.v40.1.

Alexander, C., J. Deng, J. Feng, and H. Wan. 2023. “Net Buying Pressure and the Information in Bitcoin Option Trades.” *Journal of Financial Markets* 63: 100764. https://doi.org/10.1016/j.finmar.2022.100764.

Alexander, C., and D. F. Heck. 2020. “Price Discovery in Bitcoin: The Impact of Unregulated Markets.” *Journal of Financial Stability* 50: 100776. https://doi.org/10.1016/j.jfs.2020.100776.

Aspris, A., A. H. Dyhrberg, S. Foley, and T. Putnins. 2022. Digital Assets and Markets: A Transaction-Cost Analysis of Market Architectures. Swift Institute Working Paper, 2021-001.

Aspris, A., S. Foley, J. Svec, and L. Wang. 2021. “Decentralized Exchanges: The “Wild West” of Cryptocurrency Trading.” *International Review of Financial Analysis* 77: 101845. https://doi.org/10.1016/j.irfa.2021.101845.

Bailey, W., and K. C. Chan. 1993. “Macroeconomic Influences and the Variability of the Commodity Futures Basis.” *The Journal of Finance* 48: 555–573. https://doi.org/10.1111/j.1540-6261.1993.tb04727.x.

Baker, M., and J. Wurgler. 2006. “Investor Sentiment and the Cross-Section of Stock Returns.” *Journal of Finance* 61: 1645–1680. https://doi.org/10.1111/j.1540-6261.2006.00885.x.

Barbon, A., and A. Ranaldo. 2023. On The Quality Of Cryptocurrency Markets: Centralized Versus Decentralized Exchanges. Working Paper. https://doi.org/10.2139/ssrn.3984897.

Baur, D. G., and T. Dimpfl. 2019. “Price Discovery in Bitcoin Spot or Futures?” *Journal of Futures Markets* 39: 803–817. https://doi.org/10.1002/fut.v39.7.

Brauneis, A., R. Mestel, R. Riordan, and E. Theissen. 2021. “How to Measure the Liquidity of Cryptocurrency Markets?” *Journal of Banking and Finance* 124: 106041. https://doi.org/10.1016/j.jbankfin.2020.106041.

Brunnermeier, M. K., and L. H. Pedersen. 2009. “Market Liquidity and Funding Liquidity.” *Review of Financial Studies* 22: 2201–2238. https://doi.org/10.1093/rfs/hhn098.

Cao, M., and B. Celik. 2021. “Valuation of Bitcoin Options.” *Journal of Futures Markets* 41: 1007–1026. https://doi.org/10.1002/fut.v41.7.

Cheng, I. H., A. Kirilenko, and W. Xiong. 2015. “Convective Risk Flows in Commodity Futures Markets.” *Review of Finance* 19: 1733–1781. https://doi.org/10.1093/rof/rfu043.

Chiu, J., E. Ozdenoren, K. Z. Yuan, and S. Zhang. 2023. On the Fragility of DeFi Lending. Working Paper. https://doi.org/10.2139/ssrn.4328481.

Choi, K. J., J. Jeon, and B. H. Lim. 2023. “Optimal Staking and Liquid Token Holding Decisions in Cryptocurrency Markets.” *Working Paper*: 1–2. https://doi.org/10.2139/ssrn.4528742.

Chordia, T., R. Roll, and A. Subrahmanyam. 2001. “Market Liquidity and Trading Activity.” *The Journal of Finance* 56: 501–530. https://doi.org/10.1111/jofi.2001.56.issue-2.

Coinbase Institutional. 2023. Liquid Staking: Reconciling the stETH Price Gap. https://www.coinbase.com/en-de/institutional/research-insights/research/monthly-outlook/liquid-staking-june-2022.

Cong, L., Z. He, and K. Tang. 2022. Staking, Token Pricing, and Crypto Carry. Working Paper. https://doi.org/10.2139/ssrn.4059460.

Corbet, S., B. Lucey, M. Peat, and S. Vigne. 2018. “Bitcoin Futures–What Use Are They?” *Economics Letters* 172: 23–27. https://doi.org/10.1016/j.econlet.2018.07.031.

Cornell, B., and K. R. French. 1983. “The Pricing of Stock Index Futures.” *Journal of Futures Markets* 3: 1. https://doi.org/10.1002/fut.v3:1.

Corwin, S. A., and P. Schultz. 2012. “A Simple Way to Estimate Bid-Ask Spreads From Daily High and Low Prices.” *The Journal of Finance* 67: 719–760. https://doi.org/10.1111/jofi.2011.67.issue-2.

Cretarola, A., and G. Figà-Talamanca. 2020. “Bubble Regime Identification in an Attention-Based Model for Bitcoin and Ethereum Price Dynamics.” *Economics Letters* 191: 108831. https://doi.org/10.1016/j.econlet.2019.108831.

De Blasis, R., and A. Webb. 2022. “Arbitrage, Contract Design, and Market Structure in Bitcoin Futures Markets.” *Journal of Futures Markets* 42: 492–524. https://doi.org/10.1002/fut.v42.3.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann. 1990. “Noise Trader Risk in Financial Markets.” *Journal of Political Economy* 98: 703–738. https://doi.org/10.1086/261703.

Deribit. 2023. Deribit API. https://docs.deribit.com/.

Dune Analytics. 2023. Dune API. https://dune.com/product/api.

Eichler, S. 2012. “Limited Investor Attention and the Mispricing of American Depositary Receipts.” *Economics Letters* 115: 490–492. https://doi.org/10.1016/j.econlet.2011.12.111.

Entrop, O., B. Frijns, and M. Seruset. 2020. “The Determinants of Price Discovery on Bitcoin Markets.” *Journal of Futures Markets* 40: 816–837. https://doi.org/10.1002/fut.v40.5.

Ethereum Foundation. 2022. The Risks of LSD. https://notes.ethereum.org/@djrtwo/risks-of-lsd.

FTX. 2022. FTX API [Defunct]. https://ftx.com/.

Gagnon, L., and A. Karolyi. 2010. “Multi-Market Trading and Arbitrage.” *Journal of Financial Economics* 97: 53–80. https://doi.org/10.1016/j.jfineco.2010.03.005.

Gogol, K., B. Kraner, M. Schlosser, T. Yan, C. Tessone, and B. Stiller. 2024. Empirical and Theoretical Analysis of Liquid Staking Protocols. Working Paper. https://doi.org/10.48550/arXiv.2401.16353.

Gonzalo, J., and C. Granger. 1995. “Estimation of Common Long-Memory Components in Cointegrated Systems.” *Journal of Business and Economic Statistics* 13: 27–35. https://doi.org/10.1080/07350015.1995.10524576.

Grandjean, D., L. Heimbach, and R. Wattenhofer. 2023. Ethereum Proof-of-Stake Consensus Layer: Participation and Decentralization. Working Paper. https://doi.org/10.48550/arXiv.2306.10777.

Griffin, J. M., and A. Shams. 2020. “Is Bitcoin Really Untethered?” *The Journal of Finance* 75: 1913–1964. https://doi.org/10.1111/jofi.12903.

Grobys, K., J. Junttila, J. W. Kolari, and N. Sapkota. 2021. “On the Stability of Stablecoins.” *Journal of Empirical Finance* 64: 207–223. https://doi.org/10.1016/j.jempfin.2021.09.002.

Han, J., and Z. Pan. 2017. “On the Relation Between Liquidity and the Futures-Cash Basis: Evidence From a Natural Experiment.” *Journal of Financial Markets* 36: 115–131. https://doi.org/10.1016/j.finmar.2016.12.002.

Han, L., Z. Li, and L. Yin. 2017. “The Effects of Investor Attention on Commodity Futures Markets.” *Journal of Futures Markets* 37: 1031–1049. https://doi.org/10.1002/fut.v37.10.

Han, X., N. Sakkas, J. Danbolt, and A. Eshraghi. 2022. “Persistence of Investor Sentiment and Market Mispricing.” *Financial Review* 57: 617–640. https://doi.org/10.1111/fire.v57.3.

Hasbrouck, J. 1995. “One Security, Many Markets: Determining the Contributions to Price Discovery.” *The Journal of Finance* 50: 1175–1199. https://doi.org/10.1111/jofi.1995.50.issue-4.

Hattori, T., and R. Ishida. 2021. “The Relationship Between Arbitrage in Futures and Spot Markets and Bitcoin Price Movements: Evidence From the Bitcoin Markets.” *Journal of Futures Markets* 41: 105–114. https://doi.org/10.1002/fut.22171.

Hoang, L. T., and D. G. Baur. 2021. “How Stable Are Stablecoins?” *The European Journal of Finance* 30: 1984–2000. https://doi.org/10.1080/1351847X.2021.1949369.

Hou, A. J., W. Wang, C. Y. Chen, and W. K. Härdle. 2020. “Pricing Cryptocurrency Options.” *Journal of Financial Econometrics*, 18: 250–279. https://doi.org/10.1093/jjfinec/nbaa006.

Hu, Y., Y. G. Hou, and L. Oxley. 2020. “What Role Do Futures Markets Play in Bitcoin Pricing? Causality, Cointegration and Price Discovery From a Time-Varying Perspective?” *International Review of Financial Analysis* 72: 101569. https://doi.org/10.1016/j.irfa.2020.101569.

IntoTheBlock. 2023. IntoTheBlock API. https://api.intotheblock.com/docs/index#/Cryptocurrencies.

Irresberger, F., K. John, P. C. Mueller, and F. Saleh. 2020. The Public Blockchain Ecosystem: An Empirical Analysis. Working Paper. https://doi.org/10.2139/ssrn.3592849.

Jahanshahloo, H., S. Corbet, and L. Oxley. 2022. “Seeking Sigma: Time-Of-The-Day Effects on the Bitcoin Network.” *Finance Research Letters* 49: 103101. https://doi.org/10.1016/j.frl.2022.103101.

John, K., T. Rivera, and F. Saleh. 2021. Equilibrium Staking Levels in a Proof-of-Stake Blockchain. Working Paper. https://doi.org/10.2139/ssrn.3965599.

Kadapakkam, P., and U. Kumar. 2013. “Impact of Liquidity on the Futures-Cash Basis: Evidence From the Indian Market.” *Journal of Futures Markets* 33: 266–298. https://doi.org/10.1002/fut.2013.33.issue-3.

Kapar, B., and J. Olmo. 2019. “An Analysis of Price Discovery Between Bitcoin Futures and Spot Markets.” *Economics Letters* 174: 62–64. https://doi.org/10.1016/j.econlet.2018.10.031.

Kraken. 2023. Downloadable historical OHLCVT. https://support.kraken.com/hc/en-us/articles/360047124832-Downloadable-historical-OHLCVT-Open-High-Low-Close-Volume-Trades-data.

Kurov, A. 2010. “Investor Sentiment and the Stock Market's Reaction to Monetary Policy.” *Journal of Banking & Finance* 34: 139–149. https://doi.org/10.1016/j.jbankfin.2009.07.010.

Lehar, A., and C. A. Parlour. 2021. Decentralized Exchange. Working Paper. https://doi.org/10.2139/ssrn.3905316.

Lehar, A., and C. A. Parlour. 2022. Systemic Fragility in Decentralized Markets. Working Paper. https://doi.org/10.2139/ssrn.4164833.

Lien, D., and K. Shrestha. 2009. “A New Information Share Measure.” *Journal of Futures Markets* 29: 377–395. https://doi.org/10.1002/fut.v29:4.

Lo, Y., and F. Medda. 2021. Uniswap and the Rise of the Decentralized Exchange. Working Paper. https://doi.org/10.2139/ssrn.3715398.

Milunovich, G. 2022. “Assessing the Connectedness Between Proof of Work and Proof of Stake/Other Digital Coins.” *Economics Letters* 211: 110243. https://doi.org/10.1016/j.econlet.2021.110243.

Nansen. 2022. On-Chain Forensics: Demystifying stETH's “De-peg”. https://www.nansen.ai/research/on-chain-forensics-demystifying-steth-depeg.

OECD. 2022. Lessons From the Crypto Winter: DeFi Versus CeFi. OECD Business and Finance Policy Papers. https://doi.org/10.1787/bf84ff64-en.

Roll, R., E. Schwartz, and A. Subrahmanyam. 2007. “Liquidity and the Law of One Price: The Case of the Futures-Cash Basis.” *The Journal of Finance* 62: 2201–2234. https://doi.org/10.1111/jofi.2007.62.issue-5.

Roşu, I., and F. Saleh. 2021. “Evolution of Shares in a Proof-Of-Stake Cryptocurrency.” *Management Science* 67: 661–672. https://doi.org/10.1287/mnsc.2020.3791.

Saleh, F. 2021. “Blockchain Without Waste: Proof-Of-Stake.” *The Review of Financial Studies* 34: 1156–1190. https://doi.org/10.1093/rfs/hhaa075.

Sapkota, N., and K. Grobys. 2021. “Blockchain Consensus Protocols, Energy Consumption and Cryptocurrency Prices.” *The Journal of Energy Markets* 13: 117–139. https://doi.org/10.21314/JEM.2020.221.

Scharnowski, S. 2021. “Understanding Bitcoin Liquidity.” *Finance Research Letters* 38: 101477. https://doi.org/10.1016/j.frl.2020.101477.

Schmeling, M., A. Schrimpf, and K. Todorov. 2023. Crypto Carry. Working Paper. https://doi.org/10.2139/ssrn.4268371.

Shleifer, A., and R. W. Vishny. 1997. “The Limits of Arbitrage.” *The Journal of Finance* 52: 35–55. https://doi.org/10.1111/jofi.1997.52.issue-1.

Singal, V., and J. Tayal. 2020. “Risky Short Positions and Investor Sentiment: Evidence From the Weekend Effect in Futures Markets.” *Journal of Futures Markets* 40: 479–500. https://doi.org/10.1002/fut.v40.3.

Stambaugh, R. F., J. Yu, and Y. Yuan. 2015. “Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle.” *The Journal of Finance* 70: 1903–1948. https://doi.org/10.1111/jofi.2015.70.issue-5.

Tzinas, A., and D. Zindros. 2024. “ The Principal-Agent Problem in Liquid Staking.” In Financial Cryptography and Data Security. FC 2023 International Workshops, edited by A. Essex, S. Matsuo, O. Kulyk, L. Gudgeon, A. Klages-Mundt, D. Perez, S. Werner, A. Bracciali, and G. Goodell, 456–469. https://doi.org/10.1007/978-3-031-48806-1_29.

Appendix

In this appendix, we present additional results to establish the robustness of our main findings. We first show regression results for the determinants of the basis when including the Solana-based tokens. We then provide evidence regarding the dynamic relationship between limits to arbitrage and the basis. Finally, we analyze price discovery for Solana.

Appendix A

Determinants of the Liquid Staking Bases of Individual Tokens

We first show results of panel regression in Table A1 where we include stSOL and mSOL alongside stETH as traded at the CEX. We include fixed effects for each token to control for any systematic differences in the bases of the different tokens. We furthermore include the binary variable *Crypto Winter* and interact this variable with the token fixed effects. We find that the overall conclusions do not materially differ from the findings in Table 4. Moreover, we estimate the model separately for each token. Again, the results are generally confirmed. While not all coefficients are statistically significant, those that are have the same sign as in the pooled model, except for the futures basis for stETH and the crypto winter for the Solana-based tokens.

TABLE A1 | Determinants of the liquid staking bases of individual tokens at the CEX.

	All	stETH	mSOL	stSOL
stETH × crypto winter	2.606***	2.614***		
	(129.35)	(136.29)		
mSOL × crypto winter	0.016***		−0.006***	
	(3.97)		(−3.15)	
stSOL × crypto winter	−0.000			−0.029***
	(−0.04)			(−4.53)
Weekend	0.040***	0.096***	−0.002	0.024***
	(6.13)	(6.44)	(−0.87)	(2.70)
$\text{Return}^{\text{Absolute}}_{\text{Underlying}}$	0.019***	0.021	0.012***	0.022***
	(3.72)	(1.17)	(8.11)	(4.81)
$\text{Return}^{\text{Positive}}_{\text{Underlying}}$	−0.004	0.003	0.001	−0.010*
	(−0.61)	(0.18)	(0.61)	(−1.86)
$\text{Spread}_{\text{Staking token}}$	0.832***	0.469***	1.151***	1.393***
	(11.15)	(2.65)	(19.39)	(19.25)
$\text{Volume}_{\text{Staking token}}$	0.489***	1.240***	0.131***	−0.008
	(4.01)	(4.82)	(3.44)	(−0.42)
Forced liquidations	0.029***	0.033**	0.001	0.009**
	(5.29)	(2.37)	(0.70)	(2.45)
Attention	−0.166***	−0.480***	−0.010**	−0.007
	(−11.72)	(−12.36)	(−2.34)	(−0.52)
Sentiment	1.845***	5.261***	0.216***	−0.133
	(22.20)	(30.14)	(13.12)	(−1.59)
Liquid staking basis_+	0.009**	0.062	0.013***	0.035***
	(2.27)	(1.38)	(6.69)	(7.40)
Futures basis	0.163	−2.095**	0.716***	1.172***
	(1.10)	(−2.49)	(9.62)	(7.04)
Token FE	✓	—	—	—
Hour FE	✓	✓	✓	✓
Adj. R^2	0.82	0.82	0.45	0.13
Observations	21,885	7,295	7,295	7,295
Sample end	2022-08	2022-08	2022-08	2022-08

Note: This table shows panel regression results for the determinants of the absolute liquid staking basis at the CEX. For staking tokens on Solana, the basis has been detrended using a linear time trend. Model (1) uses data on all staking tokens. In models (2)–(4), only data on stETH, mSOL, and stSOL is used, respectively. The independent variables are defined as in Table 3 and Table 2. All models include a constant and fixed effects for the hour of the day. Robust t-statistics are given in parentheses. ***, **, * denotes significance at the 1%, 5%, 10%–level, respectively.

Appendix B

Dynamic Relationship Between Basis and Limits to Arbitrage

Next, we provide additional evidence supporting Hypothesis 3 by conducting vector autoregressions (VAR) and Granger causality tests as in Roll, Schwartz, and Subrahmanyam (2007) and Kadapakkam and Kumar (2013). For brevity, we only provide the results for those variables relating to limits to arbitrage that are significant in our main regressions in Table 3 and Table 4.

In untabulated first stage regressions, we estimate adjusted values of the bases and the liquidity and volatility measures similar to Roll, Schwartz, and Subrahmanyam (2007). In particular, we individually regress these variables on binary variables indicating the hour of the day, the weekend, the Crypto Winter, and the Post Merge period. We then estimate pairwise VAR models using the residuals of the first stage regressions. We choose the lag length for each VAR individually according to the Akaike and Bayesian Information Criteria from a maximum lag length of 24. When the two criteria do not agree, we take the lower number of lags indicated by either of them. Across all estimated VARs, the average chosen lag length is about 14 (hours).

The results are presented in Table B1. Panel A shows the pairwise correlations of VAR innovations (i.e., residuals). All correlations are statistically significant except for forced liquidations, the proxy for funding liquidity. The correlation coefficients, when significant, have

TABLE B1 | Dynamic relationship between limits to arbitrage and the basis.

***Panel A:* Correlations between VAR innovations**

	$\text{Basis}^{\text{DEX}}$	$\text{Basis}^{\text{CEX}}$
Pool size	−0.279***	
Pool fees	0.133***	
$\text{Return}_{\text{Underlying}}^{\text{Absolute}}$	0.036***	0.081***
$\text{Spread}_{\text{Staking token}}$		0.337***
$\text{Volume}_{\text{Staking token}}$		0.083***
Forced liquidations		−0.002
$\text{Basis}^{\text{CEX}}$	0.303***	

***Panel B:* Granger causality tests**

	H_0: Row does not cause column		H_0: Column does not cause row	
	$\text{Basis}^{\text{DEX}}$	$\text{Basis}^{\text{CEX}}$	$\text{Basis}^{\text{DEX}}$	$\text{Basis}^{\text{CEX}}$
Pool size	74.38***		13.89***	
Pool fees	26.89**		278.81***	
$\text{Return}_{\text{Underlying}}^{\text{Absolute}}$	27.25**	43.23***	37.63***	64.35***
$\text{Spread}_{\text{Staking token}}$		212.45***		23.44**
$\text{Volume}_{\text{Staking token}}$		77.24***		67.64***
Forced liquidations		2.45		5.06
$\text{Basis}^{\text{CEX}}$	265.23***		1975.62***	

Note: This table shows results for vector autoregressions (VAR) and Granger causality tests. The variables are defined as in Table 3 and Table 2. All variables are first adjusted by taking the residuals of a regression of the variable on hour-of-day fixed effects and the binary variables *Crypto Winter* and *Post Merge*. Panel A shows correlations of innovations of bi-variate VAR models for the adjusted variables in the row and column. The lag length of the VAR is chosen according to the minimum indicated by AIC and BIC. Panel B shows chi-squared statistics of corresponding Granger causality tests. The null hypothesis is that the row (column) variable does not Granger-cause the column (row) variable for the left (right) set of columns. ***, **, * denotes significance at the 1%, 5%, 10%–level, respectively.

the same sign as in the regressions in Section 5, consistent with Hypothesis 3 regarding the co-movement of limits to arbitrage and the magnitude of the liquid staking basis.

Panel B shows the results of pairwise Granger causality tests. Most chi-squared test statistics are highly significant. In the first two columns, the null hypotheses that the row variables do not Granger-cause the column variables (the basis at the DEX and the CEX, respectively) are rejected in all cases except for forced liquidations.[21] Similarly, in the rightmost two columns, the reverse null hypotheses are rejected in all cases, again except for funding liquidity. For all these variables we hence find evidence of bidirectional Granger causality between the liquid staking basis and liquidity. Overall, these results align with those presented in Section 5 and are consistent with Hypothesis 3.

In the last row of Panel B, we consider Granger causality between the bases at the DEX and the CEX. The test statistic for the null hypothesis that the basis at the CEX does not Granger-cause the basis at the DEX is significant, but substantially smaller than for the null hypothesis in the other direction. This finding is consistent with the results of the price discovery analysis as presented in Table 6.

Finally, we analyze impulse response functions (IRFs) similar to the studies mentioned above. IRFs tracks the effect of a one-time, unit standard deviation positive shock. To orthogonalize the innovations, we use the inverse of the Cholesky decomposition of the residual covariance matrix. Figure B1 and Figure B2 show the cross responses of the absolute basis and the liquidity and volatility measures for the DEX and the CEX, respectively. The response of a variable to its own shock is omitted for brevity. Note that the reported number of lags for the IRFs vary because we select the optimum lag length according to the information criteria for each VAR individually.

The results again confirm our main findings and support Hypothesis 3. However, the IRFs shed further light on the time dynamics in the relationship between limits to arbitrage and the basis. We discuss the main observations. In the first row of Figure B1, we find that the basis is negatively impacted by positive shocks in pool size, i.e., by increases in the available liquidity at the DEX. Shocks to the basis likewise have a negative effect on pool size, but the reaction is slower. The effect of shocks to pool fees is instantaneous and gradually decreases over time. Interestingly, pool fees appear to increase gradually after a shock to the basis. Since a wider basis might indicate imbalance in the pool's assets, higher fees incentivize liquidity providers to restore balance in the pool. As shown in the last row, shocks to the basis at one exchange impact the basis at the other exchange. However, the impact of shocks to the basis at the DEX on the basis at the CEX appears stronger than the opposite direction. The results for the CEX in Figure B2 are overall similar and indicate a persistent effect of limits to arbitrage on the liquid staking basis.

Appendix C

Price Discovery

Table C1 shows the results for price discovery of Solana at the CEX. Both staking tokens contribute similar amounts to price discovery that are likewise non-trival but substantially smaller than for the spot and futures markets. Finally, Figure C1 depicts the evolution of the daily price discovery measures for Solana, again showing time-variation in the relative importance of the liquid staking derivatives. On average, their contribution is larger during the second half of the sample period.

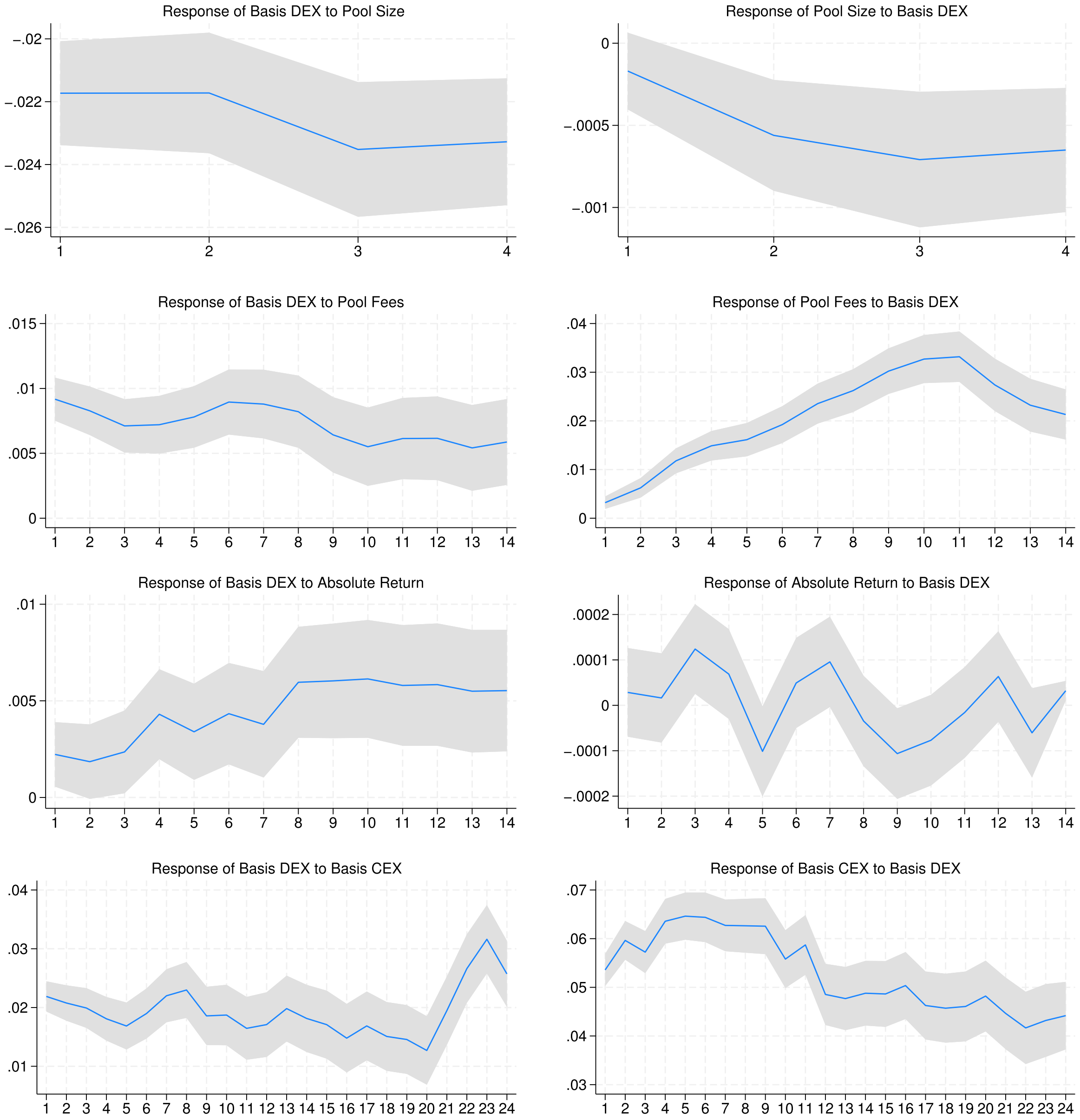


FIGURE B1 | Impulse response functions for the DEX. *Note:* These graphs show orthogonalized impulse response functions for the DEX based on bivariate vector autoregressions where the lag length is chosen according to the minimum indicated by AIC and BIC. Only cross-responses are shown. The variables are as defined as in Table 3 and Table 4 but adjusted by taking the residuals of a first-stage regression on *Crypto Winter*, *Post Merge*, *Weekend*, and the hour fixed effects.

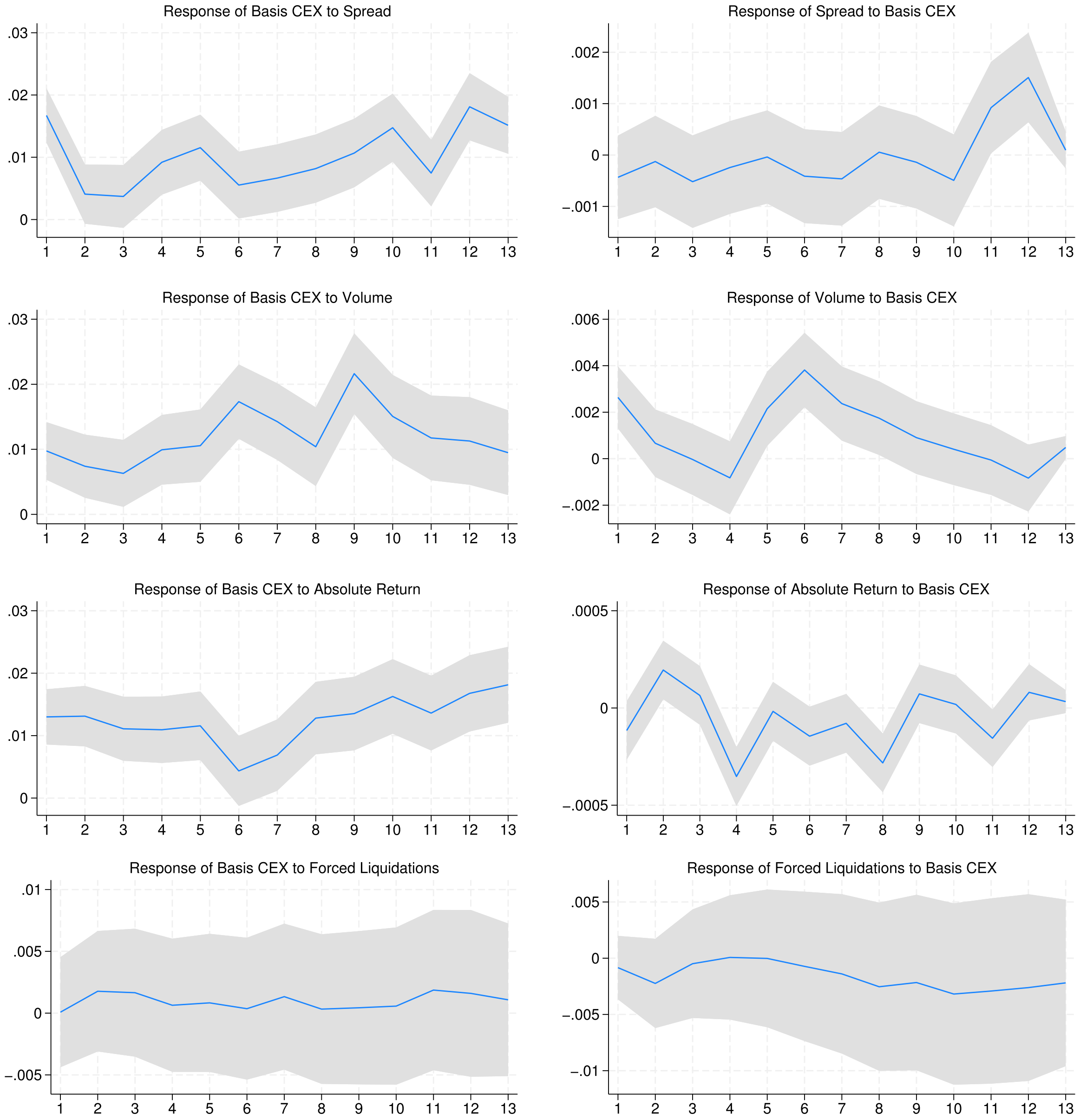


FIGURE B2 | Impulse response functions for the CEX. *Note:* These graphs show orthogonalized impulse response functions for the CEX based on bivariate vector autoregressions where the lag length is chosen according to the minimum indicated by AIC and BIC. Only cross-responses are shown. The variables are as defined as in Table 3 and Table 4 but adjusted by taking the residuals of a first-stage regression on *Crypto Winter*, *Post Merge*, *Weekend*, and the hour fixed effects.

TABLE C1 | Price discovery for Solana.

	Mod. information shares			Component shares		
	Mean	**Median**	**Overall**	**Mean**	**Median**	**Overall**
mSOL_{CEX}	4.62	3.21	1.58	6.58	5.08	0.59
$\text{stSOL}_{\text{CEX}}$	5.41	3.17	1.04	6.73	5.52	0.93
fSOL_{CEX}	46.43	47.46	52.67	45.77	49.84	65.33
SOL_{CEX}	43.54	43.33	44.72	40.92	38.60	33.14

Note: This table shows modified information shares and component shares for the Solana-based liquid staking tokens, the perpetual futures contracts (fETH and fSOL), and the spot market (SOL). Means and medians are based on daily estimates while the column *Overall* contains the estimate for the full sample. The prices of the staking tokens are adjusted to remove the mechanical price increase relative to Solana using the de-trended basis. All values are given in percent.

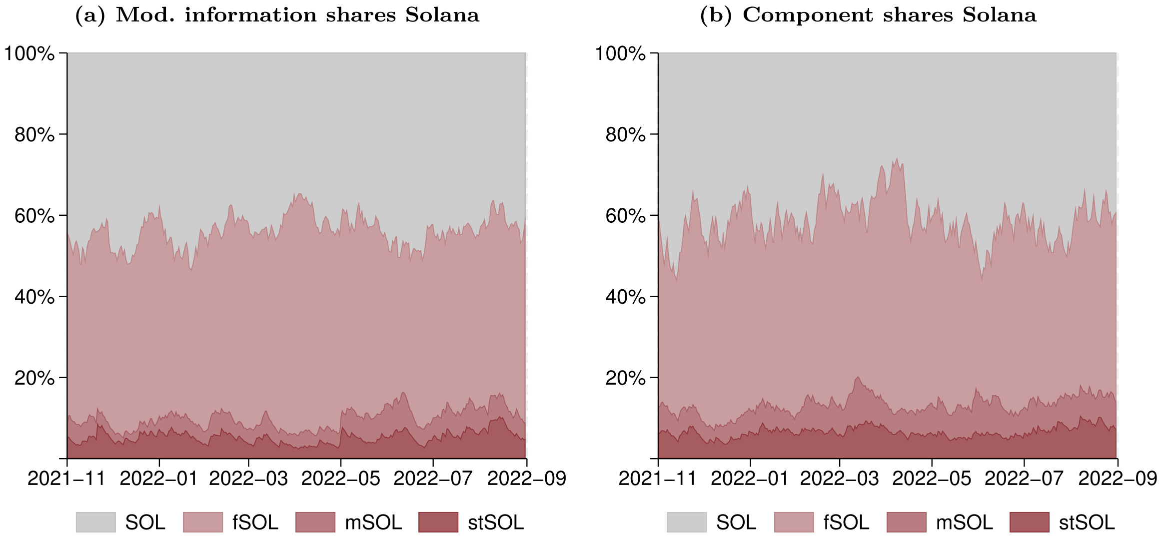


FIGURE C1 | Evolution of price discovery for Solana. *Note:* These graphs show daily modified information shares and component shares for Solana for the liquid staking tokens (mSOL and stSOL), the perpetual futures contract (fSOL), and the underlying (SOL). The prices of the Solana-based staking tokens are adjusted to remove the mechanical price increase relative to Solana using the de-trended basis. However, using unadjusted prices yields virtually identical results. For readability, the information shares are smoothed with a trailing exponentially weighted moving average using 0.1 as the smoothing parameter.