Yield Farming[*]

Patrick Augustin,[†] Roy Chen-Zhang,[‡] and Donghwa Shin[§]

First draft: October 15, 2021
This draft: September 24, 2022 (*Download the latest version here*)

Abstract

We characterize the risk and return characteristics of yield farming investment strategies on PancakeSwap, one of the largest automated market makers among the emerging ecosystem of decentralized financial services. PancakeSwap provides opportunities for earning passive income by pledging pairs of cryptocurrency tokens in liquidity pools and harvesting governance tokens in yield farms, a practice called 'yield farming.' Yield farming generates performance through several components related to capital gains, trading fee revenue and farm yields, and is exposed to impermanent losses that are driven non-linearly by differential return performance in the underlying cryptocurrency token pairs. We find that yield farming delivers positive Sharpe ratios that are comparable to those of other cryptocurrency investments and the S&P 500 index. However, investment performance declines significantly after accounting for transaction costs and price impact that is largest for farms with the highest headline yields. leading possibly to negative risk-adjusted returns. We furthermore find that flows to high-yield farms chase past performance and high yields and predict negative future returns. These patterns are similar to investment behavior and risk-return characteristics observed in traditional markets, despite the absence of financial intermediaries. Since yield farming is easily accessible to retail investors, our analysis has important implications for the current debate about the regulation of decentralized financial services.

JEL Classification Codes: G12, G13, G14, O33, Y80

Keywords: blockchain, consumer protection, cryptocurrencies, decentralized finance, fintech, household finance, yield farming

---

[*]**Preliminary Draft.** No guarantee for accuracy or sense, comments welcome. As for helpful comments, we thank Greg Brown, Mikhail Chernov, William Cong, Darrell Duffie, Christian Lundblad, Paige Ouimet, Neil Pearson, Cameron Peng (discussant), Daniel Rabetti (discussant), Adam Reed, Alessio Saretto, and participants at the 2022 Active Management Research Symposium organized by UNC Institute of Private Capital, 2022 Boulder Summer Conference, 2022 Greater China Area Finance Conference, themed "Finance Research in the Era of Big Data", 2022 International Risk Management Conference, 15th Annual Risk Management Conference at National University of Singapore, 2022 UBRI Connect at UCL, 2022 Canadian Derivatives Institute Annual Conferenece, a finance seminar at UNC Kenan-Flagler Business School. We also thank Wei Yang for excellent research assistance. Augustin acknowledges financial support from the Canadian Derivatives Institute. Shin acknowledges financial support from the Kenan Institute.

[†]McGill University and Canadian Derivatives Institute; patrick.augustin@mcgill.ca.

[‡]University of North Carolina at Chapel Hill; Roy_Chen-Zhang@kenan-flagler.unc.edu.

[§]University of North Carolina at Chapel Hill; Donghwa_Shin@kenan-flagler.unc.edu.

"Right now, we just don't have enough investor protection in crypto. Frankly, at this time, it's more like the Wild West."

Chair Gary Gensler, Securities and Exchange Commission

1 Introduction

Decentralized finance (DeFi) is a rapidly growing segment of the emerging cryptocurrency ecosystem. By operating through applications built on blockchains and executed through smart contracts, DeFi intends to eliminate the influence of central financial intermediaries. Figure 1 illustrates that the total value locked, a measure of aggregate capital investments into DeFi applications, has been growing at an exponential pace over the last 2 years, having recently surpassed $200 billion. Harvey, Ramachandran, and Santoro (2021) discuss the implications of DeFi for the future of our financial system.

We study yield farming, a novel decentralized financial service that is accessible to both retail and institutional investors. Yield farming is a modern version of securities lending whereby investors earn passive income from the provision of digital liquidity. This investment activity has gathered significant interest in the popular press because of the scale of promised returns. Advertised interest rates are often several hundred percent and can be as high as as several thousand percent (e.g., Oliver, 2021). Yield farming has also attracted regulatory interest from the Securities and Exchange Commission, who consider it to be an unregulated and complex investment strategy with hidden risks to unsophisticated investors (e.g., Gensler, 2021).

We provide the first assessment of yield farming performance using a novel hand-collected data set sourced from Binance Smart Chain (BSC). BSC is a public blockchain launched by Binance with the purpose of providing a centralized exchange that features high trade execution speeds, lower congestion risks and lower trading fees than other comparable smart contract applications. On BSC, Binance Coin (BNB) is the main medium of exchange for trading purposes and for the payment of transaction costs that are known as gas fees.

Our data contains daily information on a cross-section of 219 yield farms that are active on PancakeSwap between September 23, 2020 and September 5, 2021. PancakeSwap is one of the most popular automated market makers, next to other similar platforms like Uniswap and Sushiswap. It is also one of the largest automated market makers with 435,130 active users on October 24, 2021, compared to 47,730 active users recorded on Uniswap. PancakeSwap is particularly useful for studying yield farming since it is one of the platforms that lists yield farms in addition to liquidity pools. Studying yield farms is the centrepiece of our work. Moreover, the number of active users in relation to total value locked suggests that the yield farming space is populated by many retail investors, thereby emphasizing the importance for investor protection.

We first characterize the risk and return characteristics of yield farming strategies. Yield farming is a mechanism for passively earning income by supplying digital liquidity. The overall performance is earned through a chain of transactions that, taken together, form a complex investment product.

Yield farmers first pledge pairs of cryptocurrency tokens in equal amounts to associated liquidity pools for which they are rewarded with liquidity tokens. These liquidity tokens certify the liquidity provision and represent a claim to a fraction of the aggregate liquidity in the pool. When the liquidity tokens are redeemed, the fractional ownership of the pool may change in value due to the constant product technology hardwired into the automated market maker system. Thus, liquidity miners may face capital gains or losses, in addition to fee revenue from the trading activity of third party investors in exchange for the liquidity provision. Liquidity miners also face significant downside risk through impermanent losses. This is realized as a loss function that is driven non-linearly by return differences among the pair of cryptocurrency tokens associated with a pool.

Another key component of yield farming is the staking of liquidity tokens that are rewarded for liquidity provision into yield farms. Each liquidity pool is associated with a unique yield farm that promises an interest rate often exceeding several hundred percent. Yield farmers earn that rate in proportion to the aggregate liquidity locked in a yield farm, which is paid using PancakeSwap's governance token Cake. This ultimately exposes yield farmers further to exchange rate risk stemming from the variation in the value of Cake relative to the USD.

We explicitly show that yield farming is subject to non-trivial price impacts that are non-linearly related to the amount of liquidity invested into a yield farming strategy. We formalize the price impact function and show that it is concave in the investment amount. Besides the price impact, we also provide information on the magnitude and time variation of gas fees charged for round trip costs in yield farming.

As a second step, we assess the empirical return performance of yield farming strategies and compare them to other benchmark strategies in cryptocurrency markets and the S&P500 index. We take the perspective of U.S. investor who needs to buy digital assets using the USD as a base currency and exchange all farm yields back to its local currency at the prevailing exchange rates. We find that yield farming strategies appear to generate attractive returns, with Sharpe ratios between 2 and 3. While such Sharpe ratios appear extraordinarily large, they are similar to those for investments into the S&P500 index, Bitcoin or Ethereum, and are partially explained by the extraordinary bull run in most asset markets during our sample period.

Yield farming also generates Sharpe ratios that are larger than those of simple buy-and-hold trading strategies in the underlying pairs of cryptocurrency tokens. It also generates superior performance to a strategy that considers liquidity mining without yield farming. Even though the joint investment activity is a strictly dominating strategy, not all investors appear to stake their liquidity tokens into yield farms. This is suggestive evidence of investor inertia and lack of investor sophistication.

While we uncover positive investment performance without the consideration of transaction costs, the performance becomes significantly weaker when we account for trading costs (a.k.a. gas fees) and price impact. While gas fees shift the return performance linearly downward for all yield farms, price impact is especially important for farms that advertise large headline yields. This motivates our additional analysis on the relation between flows and performance in yield farms.

As a last step, we study the relation between yield farming flows and performance. We follow the mutual fund literature and define farm flows as the change in total value locked, after accounting for growth in liquidity associated with return performance. We scale flows by total value locked to make them comparable across yield farms. We find that farms with high headline yields attract more flows and that positive return performance predicts future flows. Moreover, we find that new flows are negatively correlated with future farm performance.

Overall, our evidence is consistent with evidence from other asset markets that reflects return chasing behavior, whereby flows chase positive past performance. Our findings also provide supportive evidence for patterns that are associated with reaching for yield. We consider these findings to be intriguing since they typically arise in a setting with financial intermediaries, while yield farming is implemented in a decentralized market without financial intermediaries.

Given the risks to retail investors associated with obfuscated investment strategies and complex financial products, we believe that our evidence has important policy implications for regulatory disclosure and investor protection.

2 Literature

Our work is most closely related to the emerging literature on decentralized finance. In Table 1, we compare a select number of studies that examine decentralized exchanges(DEX). To our knowledge, this is the first empirical study of the risk and return characteristics of yield farming strategies based on a novel hand collected data set.

In contrast to earlier studies, we focus on PancakeSwap, a smart contract platform operating on the Binance Smart chain (BSC). BSC charges comparatively lower transaction costs than Ethereum, the supporting blockchain for Uniswap and SushiSwap. Hence, yield farms operating on BSC are more easily accessible to retail investors.

Several studies characterize the theoretical properties of automated market makers (AMM) with the constant product model that has been adopted by major decentralized exchanges (Angeris, Kao, Chiang, Noyes, and Chitra, 2019; Aoyagi, 2021). Lehar and Parlour (2021), Park (2021) further examine strategic trading and liquidity provision in decentralized exchanges. Capponi and Jia (2021) underscore the investor trade-offs arising from the personal benefits of token investments and loss exposure associated with high token exchange rate volatility. Relatedly, we explicitly characterize the impermanent-loss and price-impact functions implicit in liquidity mining and yield farming. Han, Huang, and Zhong (2021) suggest that trading on DEXs are informative about the decentralized consensus of cryptocurrency value. Li and Mayer (2021) study the safe asset properties of stablecoins.

Yield farming is a complex and opaque investment strategy. Thus, we also relate to the literature on complex structured finance. For example, Henderson and Pearson (2011) suggest that highly popular structured retail products (SRPs) deliver subpar performance for retail investors in spite of high promised returns. Supply-based theories explain the popularity of SRPs among retail investors by arguing that intermediaries exploit investors' lack of financial sophistication (e.g. Célérier and Vallée, 2017; Egan, 2019; Ghent, Torous, and Valkanov, 2019; Henderson, Pearson, and Wang, 2020). Shin (2021) advocates a demand-based explanation whereby investors extrapolate and aggressively chase past performance. In a significant departure from that work, we study complex financial products offered through smart contracts operating on a blockchain without centralized financial intermediaries who may drive the security design or benefit from sales.

Yield farms promise passive income at impressive headline rates. This connects our work to the literature on "reaching for yield," i.e., investors' propensity to buy riskier assets to achieve higher yields. That phenomenon has been documented in the corporate bond (Becker and Ivashina, 2015; Chen and Choi, 2021) and mutual fund market (Choi and Kronlund, 2018). Bordalo, Gennaioli, and Shleifer (2016) show how investors' salience bias can lead to reaching-for-yield behavior when firms compete for consumer attention. Our evidence suggests that reaching for yield may also exist in decentralized exchanges even in the absence of financial intermediaries and related agency conflicts.

Our work adds to the developing literature of cryptocurrencies and blockchain technologies (Harvey, 2016; Yermack, 2017; Biais, Bisiere, Bouvard, and Casamatta, 2019; Saleh, 2021; Easley, O'Hara, and Basu, 2019), including initial coin offerings (e.g., Howell, Niessner, and Yermack, 2020; Hu, Parlour, and Rajan, 2019; Lee, Li, and Shin, 2022), price manipulations(Gandal, Hamrick, Moore, and Oberman, 2018; Griffin and Shams, 2020; Cong, Li, Tang, and Yang, 2021; Li, Shin, and Wang, 2021) and illegal activity (Foley, Karlsen, and Putnins, 2019), equilibrium pricing of bitcoin (Biais, Bisiere, Bouvard, Casamatta, and Menkveld, 2020; Pagnotta and Buraschi, 2018) and its adoption (Hinzen, John, and Saleh, 2020), and cryptocurrency valuation (Cong and He, 2019; Cong, Li, and Wang, 2021; Sockin and Xiong, 2020).

Makarov and Schoar (2019, 2020) document arbitrage opportunities across centralized cryptocurrency exchanges. The apparent price dispersions have been related to explanations including noise traders (Krückeberg and Scholz, 2020; Dyhrberg, 2020), liquidity frictions (Kroeger and Sarkar, 2017), settlement latency (Hautsch, Scheuch, and Voigt, 2019), risk premiums (Borri and Shaknov, 2021), restrictions to cross-border capital flows (Yu and Zhang, 2018; Choi, Lehar, and Stauffer, 2018). Augustin, Rubtsov, and Shin (2021) document an increase in bitcoin's price efficiency following the introduction of bitcoin futures.

3 Institutional background

We first provide background information on decentralized finance, yield farming, the Binance Smart Chain, and PancakeSwap. We then discuss why PancakeSwap is especially useful for the study of yield farming.

3.1 Decentralized finance and cryptocurrency yield farming

Decentralized finance (DeFi) corresponds to an emerging ecosystem of protocols and financial applications built on blockchain technology with programmable capacities, such as Ethereum and Binance Smart Chain. Smart contracts on the blockchain execute all transactions automatically, without third-party intervention.

According to DeFi Llama[1], a public dashboard which provides data on DeFi, the total dollar value locked (TVL) in decentralized financial services is $205.76 billion as of October 11, 2021. This represents a dramatic increase from less than $1 billion in February 2020.

Yield farming is a way of earning income as compensation for providing liquidity to liquidity pools. Holders of cryptocurrency tokens earn rewards by locking tokens up in liquidity pools, which issue claims to the pledged tokens. These new claims, called 'LP tokens' or 'flip tokens', can be pledged to yield farms that promise yield enhancements. That additional passive income is paid to yield farming investors using the platform's governance token.

To an extent, yield farming is a decentralized variant of securities lending, although the chain of transactions is more complex. The main reason underlying its popularity is the critical need for platform owners to incentivize liquidity provision to ensure a platform's long-term success. In a decentralized exchange, a more liquid pool implies a smaller price impact per trade, which is desirable for traders. In a lending pool, a larger amount of liquidity in a pool may drive down the borrowing interest rate, which could attract larger groups of borrowers. Yield farming is a useful tool to encourage the injection of such liquidity.

Headline rates and promised investment rates in yield farms can be large. Annual yields north of 100% are commonly observed. There exists, however, significant cross-sectional heterogeneity in promised yields across the farms, as we show in Figure 2.

Yield farming strategies appear to be complex. The total return performance from yield farming has four components: the realized yield, capital gains from pairs of cryptocurrencies, fees from trading volume in liquidity pools and yield farms, and impermanent losses driven by the relative price change of cryptocurrency pairs locked in liquidity pools. Thus, the complexity of yield farming strategies resembles obfuscated investment strategies observed

[1]https://defillama.com/home. See also Figure 1.

in complex structured derivative products (e.g. Henderson and Pearson, 2011; Célérier and Vallée, 2017; Egan, 2019; Henderson, Pearson, and Wang, 2020; Shin, 2021).

We focus our analysis on yield farms listed on PancakeSwap, a popular automated market maker that ranks second in the league tables of decentralized exchanges offering cryptocurrency lending services. Transaction costs in PancakeSwap are significantly lower than in other popular decentralized exchanges like Uniswap. This lowers the barriers to entry for retail investors, who are active investors in yield farms .

The combination of low barriers to entry, a large number of service providers, and complex investment strategies promising high returns with significant downside risk raises concerns about the protection of retail investors in cryptocurrency markets. These concerns are underscored by the aggressive stance recently taken by the U.S. Securities and Exchange Commission, who have become increasingly vocal about enhanced regulatory scrutiny of decentralized financial services. Our work is intended to inform this ongoing debate by means of assessing the risk and return characteristics of yield farming strategies.

3.2 Binance Smart Chain

Binance Chain was launched by Binance in April 2019. Its main goal is to allow for faster decentralized trading. The largest and most well-known decentralized application on the Binance Chain is Binance DEX. Despite its success in DEX trading, Binance DEX embeds several limitations that limit its flexibility. For example, to guarantee high throughput, the application does not support smart contracts, which require excess computational resources. This can, therefore, easily congest the entire network.

Binance Smart Chain (BSC) is a public blockchain running in parallel to the Binance Chain. Distinctive features of BSC include smart contract functionality and compatibility with the Ethereum Virtual Machine (EVM). BSC was launched with the purpose of maintaining the high throughput of Binance Chain while still allowing for smart contracts within the ecosystem.

In the BSC ecosystem, Binance Coin (BNB) is used as the basic medium of exchange, similar to the role played by Ether (ETH) in the Ethereum network. End users pay their transaction fees in BNB and use BNB to trade cryptocurrencies on the many decentralized exchanges deployed on BSC.

The primary advantages of BSC are its high throughput rate and low transaction fees. BSC updates its blocks approximately every 3 seconds, using a variant of the Proof-of-Stake consensus algorithm. More specifically, it employs Proof-of-Staked Authority (or PoSA), in which participants stake BNB to become validators of the blocks. As of September 5, 2021, the platform's 21 active validators play an important role in keeping the network running.

According to the CEO of Binance, Changpeng Zhao[2], BSC allows for a maximum of 300 transactions per second. In contrast, Ethereum processes up to a maximum of 16 transactions per second. The current version of BSC is thus about 20 times faster than Ethereum.

BSC transaction fees are also cheaper than those of Ethereum. As of September 5, 2021, the average transaction fee charged by BSC is $0.399, whereas the average transaction fee charged by Ethereum is $5.842. In fact, the difference in fees widens significantly when the Ethereum network becomes congested. For example, the average Ethereum transaction fee was $71.72 on May 19, 2021, whereas the maxium daily average transaction fee of BSC was $1.08 on May 11, 2021.[3]

These advantages make BSC one of the strongest competitors to Ethereum. As of October 9, 2021, total transactions on BSC have outpaced those on Ethereum, despite Ethereum preceding BSC by almost 4 years.[4] Binance Coin is currently the third largest cryptocurrency in terms of market capitalization, following Bitcoin and Ethereum.

Another important feature of the BSC is its EVM-compatibility. This implies that the chain can benefit from the rich universe of Ethereum tools and DApps. For example, project developers can easily transition their projects between Ethereum and BSC. The growth of PancakeSwap is in part spurred by the popularity of Uniswap, which is built on the Ethereum blockchain. This is because a significant part of Uniswap's source code was directly ported to BSC to build an initial version of PancakeSwap.

3.3 PancakeSwap

PancakeSwap is the largest decentralized exchange built on the Binance Smart Chain. Unlike traditional financial markets employing market-maker systems based on limit order books, PancakeSwap employs a new system called automated market maker (AMM), implemented through smart contracts. For details on the mechanism of AMMs and their pricing schedules, see, for example, Lehar and Parlour (2021).

In PancakeSwap, multiple liquidity pools are deployed to facilitate trading of pairs of cryptocurrencies. Investors deposit an equal dollar amount of two cryptocurrencies into a liquidity pool, and thereby become liquidity providers. In exchange for the liquidity provision, the liquidity provider receives LP tokens to certify their liquidity provision. In return for their liquidity provision, liquidity providers receive a fixed proportion of trading volume registered in a pool. Third-party trades on PancakeSwap are charged a fee proportional to 0.25% of the trading volume, of which 0.17% is added to the liquidity pool associated with the corresponding cryptocurrency pair.

[2] https://twitter.com/cz_binance/status/1361596039698944000.

[3] https://ycharts.com/indicators/ethereum_average_transaction_fee and https://ycharts.com/indicators/binance_smart_chain_average_transaction_fee_es

[4] Ethereum launched on July 2015, whereas Binance Smart Chain launched on April 2019.

In addition to the income generated from trading fees, liquidity providers can earn additional passive income if the liquidity pool has a corresponding yield farm. Such income, called farm yield, is earned by staking the LP tokens to the corresponding yield farm. Farm yields are paid in PancakeSwap's governance token.

PancakeSwap migrated from version 1 (v1) to version 2 (v2) on April 24, 2021. This transition was implemented to enhance the platform's technological and security features. Both versions have co-existed since then. We study yield farming for both versions.

In PancakeSwap, the CAKE token serves as the governance token for the Decentralized Autonomous Organization (DAO), where token holders can cast votes to influence the future development of the platform.

3.4 PancakeSwap as an ideal laboratory to study yield farming

Numerous decentralized trading venues offer passive income opportunities through yield farming. Among many DeFi platforms, Uniswap and PancakeSwap consistently lead the league ranks in terms of their trading activity. The key difference between these two platforms is that Uniswap runs on the Ethereum blockchain, while PancakeSwap runs on the Binance Smart Chain.

Several features of PancakeSwap make it particularly appealing for the study of yield farming. First, and most importantly, Uniswap does not offer yield farms: Liquidity providers in Uniswap liquidity protocols receive a fixed fraction of trading volume as their passive income. However, liquidity providers cannot stake their LP tokens in farms in Uniswap to earn additional income through yield farming.

Second, PancakeSwap is one of the largest decentralized exchanges. In Table 2, we report the daily trading volume for the ten largest decentralized exchanges as of October 9, 2021. The largest DEX is dYdX, which is specialized in derivatives trading. Augustin, Rubtsov, and Shin (2021) discuss the market for regulated and unregulated cryptocurrency derivatives.

The second largest DEX is PancakeSwap (v2) with a 24-hour trading volume of $1,185.34 on October 9, 2021. PancakeSwap (v2) is followed by Uniswap (v3), 1inch Liquidity Protocol, Uniswap (v2), and SushiSwap. The trading volume on PancakeSwap (v2) is comparable to the combined trading volumes of Uniswap (v3) and Uniswap (v2). While the rank tables vary over time, PancakeSwap is among the leading DEXs focused on spot trading.

Third, the low transaction cost and high transaction speed of Binance Smart Chain make PancakeSwap easily accessible to retail investors. As discussed in Section 3.2, transaction costs of the Binance Smart Chain are an order of magnitude lower than those of Ethereum. Yet, the transaction speed of Binance Smart Chain is faster than that of Ethereum. According to DappRadar[5], PancakeSwap registered 435,130 active users on October 24, 2021,

[5]DappRadar: https://dappradar.com/rankings

in contrast to 47,730 active users recorded for Uniswap. The number of active users is highest for PancakeSwap among all decentralized applications built on all blockchains tracked by DappRadar. In light of the growing concern about the risks of complex yield farming strategies for retail investors, our study has policy implications for investor protection.

Fourth, PancakeSwap features a large cross-section of yield farms. This provides important variation to help understand the risk and return characteristics of yield farms. We study 219 unique yield farms created as of September 5, 2021.

4 Conceptual framework

Yield farming enables investors to earn passive income off their cryptocurrency holdings by making tokens available for trading in decentralized markets called farms. Conceptually, this is akin to a modern version of securities lending with the distinctive feature that smart contracts operating on permissionless blockchains automatically execute transactions without involvement of financial intermediaries. In practice, the yield farming process is more complex than traditional securities lending and involves a chain of transactions that, linked together, deliver returns from "farming for yields."

Formally, yield farming involves two independent investment decisions. First, an investor can earn passive income by providing liquidity to a liquidity pool. There is a large menu of liquidity pools that operate within a decentralized trading platform like PancakeSwap. Liquidity providers receive liquidity tokens (a.k.a. LP tokens or flip tokens) in exchange for their liquidity provision.

Second investors can stake LP tokens into a yield farm to earn additional passive income. That income is paid in the form of the cryptocurrency token that makes up the base currency of the initial liquidity pool.

The total yield farming return between day t and $t+h$, $R_{t,t+h}$, is thus equal to:

$$R_{t,t+h} = R^{\ell}_{t,t+h} + R^{f}_{t,t+h}, \tag{1}$$

where $R^{\ell}_{t,t+h}$ and $R^{f}_{t,t+h}$ define the returns from liquidity provision and the staking of LP tokens into a yield farm, respectively.

4.1 Liquidity Provision

Liquidity pools are defined in terms of pairs of cryptocurrency tokens. For example, one of the most popular liquidity pools on PancakeSwap is an automated market for buying and selling ETH and the cryptocurrency token BNB. To provide liquidity to the ETH/BNB pool, a liquidity provider needs to deposit ETH and BNB in equal amounts, taking into account their current market prices. For example, if the price of one ETH corresponds to 10 BNB, an investor would need to deposit 10 BNBs for each unit of ETH.

The pools' aggregate liquidity L_t is characterized by the aggregate token valuation, defined by the number of ETH and BNB tokens, α_t^A and α_t^B, and their prices, P_t^A and P_t^B, respectively:

$$L_t = \alpha_t^A \cdot P_t^A + \alpha_t^B \cdot P_t^B. \tag{2}$$

Returns to liquidity provision are derived from two sources: growth in the value of the liquidity pool and fee revenue earned from third party trading activity in the pool, that is:

$$\begin{aligned} R_{t,t+h}^{\ell} &= \frac{L_{t+h}}{L_t} + Trading\ Fee\ Return_{t,t+h} \\ &= \frac{\alpha_{t+h}^A \cdot P_{t+h}^A + \alpha_{t+h}^B \cdot P_{t+h}^B}{\alpha_t^A \cdot P_t^A + \alpha_t^B \cdot P_t^B} + Trading\ Fee\ Return_{t,t+h}. \end{aligned} \tag{3}$$

Intuitively, growth in the value of the liquidity pool is similar to a traditional price return. The key difference is that the number of shares α_t^i is neither constant nor based on the initial investment. Instead, it is time-varying and determined by the trading activity in the liquidity pool. This feature arises because of the constant-product technology hardwired into liquidity pools. See Lehar and Parlour (2021) for details.

In exchange for their liquidity provision, investors receive LP tokens to certify their partial ownership in the pool. While the fractional ownership stays constant over time, the pool's liquidity value may change when end users independently buy and sell ETH and BNB. The terms of trade for end users are such that the product of the quantities available in the pool is equal to a constant k:

$$k = \alpha_t^A \alpha_t^B = \alpha_{t+h}^A \alpha_{t+h}^B. \tag{4}$$

This implies that the fractional claim to the liquidity pool is constant over time. However, the number of units of ETH and BNB represented by this claim will change as a result of variation in the pool's composition arising from trading activity by end users. Thus, when a liquidity provider decides to redeem their liquidity tokens in exchange for ETH and BNB, the number of tokens they receive from redemption may differ from those initially deposited (i.e., $\alpha_{t+1}^i \neq \alpha_t^i$) despite the same fractional claim to the liquidity pool.

A second feature of the constant-product technology is that the products of price and quantity have to equalize across assets, that is, for all t:

$$\alpha_t^A P_t^A = \alpha_t^B P_t^B. \tag{5}$$

A consequence of the constant-product technology is that the returns to liquidity provision have two distinct components. Investors are exposed to capital gains/losses resulting from joint changes in the tokens' prices and in the pool's liquidity, since this leads to fluctuations in the composition of tokens that an investor can claim using the liquidity token. In addition, investors are exposed to impermanent losses, which depend on the relative returns of both ETH and BNB (i.e., changes in the ratio of token prices). To formalize our discussion, the return from liquidity growth can be expressed as:

$$\frac{L_{t+h}}{L_t} = \underbrace{\left(\frac{1}{2}R^A_{t,t+h} + \frac{1}{2}R^B_{t,t+h}\right)}_{\text{capital gain}} - \underbrace{\frac{1}{2}\left(\sqrt{R^A_{t,t+h}} - \sqrt{R^B_{t,t+h}}\right)^2}_{\text{impermanent loss}}, \tag{6}$$

where $R^A_{t,t+h} = P^A_{t+h}/P^A_t$ and $R^B_{t,t+h} = P^B_{t+h}/P^B_t$ denote the gross returns of tokens A and B, corresponding to ETH and BNB in our example. In Appendix B.1, we explicitly show how the above expression is obtained from the initial liquidity provision that starts with a nominal dollar investment.

Intuitively, the impermanent loss corresponds to the difference between the return from liquidity provision and the return from a buy-and-hold strategy (without pledging the cryptocurrency tokens to a liquidity pool). Impermanent losses depend non-linearly on the relative difference in token returns. Importantly, they are strictly negative and expose investors to significant upside and downside risk analogous to a short volatility exposure (Aigner and Dhaliwal, 2021). See Appendix B.1 for additional discussion.

The total return from liquidity provision may nonetheless exceed that of a simple buy-and-hold strategy due to the additional income generated from trading fees. As of August 14, 2021, PancakeSwap charges a trading cost equivalent to 25 basis points (bp) of trading volume. Part of that (17bp) is passed on to liquidity providers as a fraction c of total trading volume $V_{t,t+h}$ observed over two consecutive time periods t and $t+h$ and proportional to the initial fractional dollar investment I_t/L_t in the liquidity pool. Since the return from trading fees depends on the initial investment, the total fee return is characterized as

$$\mathit{Trading\ Fee\ Return}_{t,t+h} = c \cdot \left((I_t/L_t)V_{t,t+h}\right)/I_t = c \cdot V_{t,t+h}/L_t. \tag{7}$$

4.2 Yield farming

A second passive source of income is generated by staking the liquidity tokens in yield farms which promise a yield y_t. That income is paid in terms of the platforms's governance token, which corresponds to Cake in the case of PancakeSwap.

The annualized yield is implicitly defined through a complicated function that depends on (a) the number of Cake tokens created through the validation of a new block on the blockchain; (b) the total number of Cake tokens redistributed for staking M_t; (c) a farm-specific multiplier m_t which defines the number of Cake tokens allocated to the farm with the creation of a new block; (d) the total liquidity staked to the farm L^{staked}_t; and (e) the price of Cake P^{Cake}_t.

The creation of new Cake tokens through blockchain validation corresponds to a rate of approximately 40 Cake tokens for each three second period. Thus, assuming that 28,800 blocks are created each day, the annualized promised yield from staking liquidity tokens to a yield farm is given by:

$$y_t = \left(\frac{365 \times 28,800 \times 40 \times m_t}{M_t}\right)\left(\frac{P_t^{Cake}}{L_t^{staked}}\right). \tag{8}$$

Cake tokens may be allocated to other purposes than yield farming. Therefore, the aggregate multiplier does not have to correspond to the sum of all multipliers across yield farms on a platform like PancakeSwap, i.e., $M \neq \sum_k m^k$, where k corresponds to the number of farms. Note that we explicitly write out the price of the Cake reward for yield farming, P_t^{Cake}, because one of the token pair in the liquidity pools does not have to be Cake. Realized farm yield between t and $t+h$ is thus defined as

$$P_{t+h}^{Cake}\sum_{n=1}^{h}\left(\frac{y_{t+n-1}}{P_{t+n-1}^{Cake}}\right)\left(\frac{1}{365}\right). \tag{9}$$

4.3 Aggregation

Aggregating across all components allows us to decompose the total $(h-period)$ return to yield farming strategies into four components associated with token capital gains, impermanent losses, revenues from trading fees, and realized farm yields:

$$\begin{aligned} R_{t,t+h} &= \underbrace{\left(\frac{1}{2}R_{t,t+h}^{A}+\frac{1}{2}R_{t,t+h}^{B}\right)}_{\text{capital gain}} - \underbrace{\frac{1}{2}\left(\sqrt{R_{t,t+h}^{A}}-\sqrt{R_{t,t+h}^{B}}\right)^2}_{\text{impermanent loss}} \\ &+ \underbrace{c \cdot V_{t,t+h}/L_t}_{\text{trading fee revenue}} + \underbrace{P_{t+h}^{Cake}\sum_{n=1}^{h}\left(\frac{y_{t+n-1}}{P_{t+n-1}^{Cake}}\right)\left(\frac{1}{365}\right)}_{\text{realized farm yield}}. \end{aligned} \tag{10}$$

4.4 Impact of trading frictions

In practice, yield farming involves a chain of transactions that, taken together, may involve sizable transaction costs. Table A.1 breaks down the chain of transactions for a hypothetical yield farming strategy. We provide additional details in Appendix B.2.

Harvesting yields at PancakeSwap involves a chain of 12 transactions (excluding step 1 and 14 in Table A.1 that are unrelated to the yield farmer's transactions). A full round-trip transaction involves three types of costs associated with gas fees, trading fees, and price impact. These costs may significantly lower the returns from yield farming.

Gas fees correspond to transaction costs associated with the use of BSC's computational resources for trade execution. Among the set of 12 transactions, yield farmers have to pay gas fees for 10 of them. The average gas fee for a round-trip of an yield farming in PancakeSwap is estimated to be \$3.28 in our sample period.

Gas fees are especially detrimental to smaller retail investors since the flat fee is more costly for small stake investments and frequent rebalancing. In addition, since the gas fee applies to each yield farm, it reduces the benefits of diversifying systematic risk across several yield farms. An initial \$1,000 investment will thus lose about 33 bps in a round-trip transaction due to gas fees alone, and 33 bps per week for weekly rebalancing. That consideration is important for retail investors who have a tendency to rebalance too frequently Odean (1999). A diversification strategy across 10 farms would incur a per period cost of 10×3.28 = \$32.8, which, for a \$1,000 investment, is more than the typical performance fee owed to a hedge fund, excluding any consideration for hurdle fees or water marks.

Gas fees thus encourage larger and more concentrated investments, which may not be appropriate for financially unsophisticated investors. In our analysis, we consider investment sizes of \$5,000, \$10,000, \$100,000 and \$1,000,000. This allows us to consider cases where gas fees do not wash out all potential yield farm returns.

Investors also incur trading fees. PancakeSwap charges a fee of 0.25% (proportional to trading volume) for each transaction. Since yield farmers need to buy and sell tokens in intermediate steps, they will lose at least an additional 0.50% of their initial investment for a round-trip transaction. See Appendix B.2 for more details.

The third transaction cost arises through price impact. To quantify price impact, we assume that yield farmers invest an amount I_t corresponding to a constant fraction f of the liquidity pool value L_t, i.e. $I_t = f \cdot L_t$. Equation (6) provides the return to liquidity provision without frictions. With price impact and ignoring trading fees, the return to liquidity provision is impacted as follows:

$$\lambda(f)\left[\left(\frac{1}{2}R^A_{t,t+h}+\frac{1}{2}R^B_{t,t+h}\right)-\frac{1}{2}\left(\sqrt{R^A_{t,t+h}}-\sqrt{R^B_{t,t+h}}\right)^2\right], \tag{11}$$

where $\lambda(f)$ is the price impact function. We illustrate in Panels (a) to (c) of Figure 3 how price impact relates to investment size. Considering both trading fees and price impact, the return to liquidity provision reduces to:

$$(1-0.0050)\lambda(f)\left[\left(\frac{1}{2}R^A_{t,t+h}+\frac{1}{2}R^B_{t,t+h}\right)-\frac{1}{2}\left(\sqrt{R^A_{t,t+h}}-\sqrt{R^B_{t,t+h}}\right)^2\right].$$

We emphasize another indirect channel through which yield farming performance is negatively affected. Equation (8) suggests a negative relation between the aggregate liquidity in a yield farm and the offered farm yield. We provide empirical support for that pattern in Figure 4. Since liquidity provision increases the size of a farm, it mechanically decreases the offered farm yield. Hence, too much liquidity provision can be a self-defeating strategy.

4.5 Yield farm flows

In our analysis, we examine flows into yield farms. To measure net inflows of liquidity, we, therefore, follow the mutual fund literature (e.g., Sirri and Tufano, 1998; Coval and Stafford, 2007) and define the measure $flow_{t,t+h}$ over an h-period trading horizon

$$Flow_{t,t+h} = \frac{L_{t+h} - L_t \times R^*_{t,t+h}}{L_t}, \tag{12}$$

where $R^*_{t,t+h}$ corresponds to the yield farm return defined in Equation (10) net of the realized farm yield, that is $R^*_{t,t+h} = \left(\frac{1}{2}R^A_{t,t+h} + \frac{1}{2}R^B_{t,t+h}\right) - \frac{1}{2}\left(\sqrt{R^A_{t,t+h}} - \sqrt{R^B_{t,t+h}}\right)^2 + c \cdot V_{t,t+h}/L_t$. We exclude the realized farm yield term in our flow definition because it does not affect the size of next period's liquidity pool, unlike capital gains, impermanent losses and trading fees. The reason is that the farm yield is paid in Cake rather than using the base cryptocurrency of th liquidity pool or yield farm.

5 Data

We assemble a novel data set of historical yield farm data, including token prices, liquidity staked to liquidity pools/yield farms and the corresponding token shares, the time-varying value of liquidity tokens that represent claims to token shares, as well as the yield farm multipliers. We obtain these data by tracing all transactions on the Binance smart chain, which forms the underlying plumbing of the PancakeSwap trading platform. We combine the yield farm data with price and volume data from Binance Smart Chain and cryptocurrency return factors constructed as in Liu, Tsyvinski, and Wu (2019).

5.1 Farm data

We collect information on all yield farms stored in PancakeSwap's main staking contract from the beginning of its existence on September 23, 2020 to September 5, 2021. We first extract from the main staking contract the contract addresses of all liquidity pools that have a yield farm associated with them. We then reconstruct from the blockchain the time series of each yield farm's multiplier at a daily frequency.

An important consideration in our study is the migration from PancakeSwap v1 to PancakeSwap v2 on April 24, 2021, when the platform upgraded the technological and security features of its smart contract design. At the time of the switch, liquidity providers in version 1 were encouraged to withdraw their liquidity and redeposit it in the corresponding liquidity pool that was transitioned to version 2. Migrating one's liquidity was considered to be a dominant strategy since the new version would pay higher rewards, and failure to migrate could lead to higher transaction costs. Both versions have co-existed since then. Interestingly, we observe that not all liquidity providers have migrated to the new version. Since each liquidity pool is uniquely linked to one pair of cryptocurrency tokens, we could easily identify the contract address of each matching liquidity pool in PancakeSwap v2.

To measure a farm's yield, we use information on the cryptocurrency shares provided to each liquidity pool (α_t^i), the aggregate liquidity in each pool (L_t), and the total amount of liquidity tokens staked in each farm (L_t^{staked}). We collect information on cryptocurrency shares using the tokens' balances in each pool. Given the tokens' prices, a pool's aggregate liquidity is computed as the aggregate dollar value of a token pair, say, $L_t = P_t^A \alpha_t^A + P_t^B \alpha_t^B$.

We further collect each pool's aggregate supply of liquidity tokens and the number of such tokens staked to the yield farms. The aggregate liquidity staked to a farm is then given by a pool's aggregate liquidity times the fraction of liquidity tokens that have been staked, $L_t^{staked} = (\#\text{ staked LP tokens}/\text{Aggregate }\#\text{ of LP tokens}) \cdot L_t$.

Given the novelty of our yield farm data and the lack of reliable information providers, we implement several accuracy checks to vet the data reliability. The most important input to our study is the farm yield described in Equation (8). We verified its accuracy by collecting offered farm yields from PancakeSwap's homepage[6] at midnight Greenwich Meridian Time (GMT) on October 11, 2021. We manually verified that the multipliers collected from the main staking contract are identical to those advertised on PancakeSwap's web interface. Then, we verified that our imputed farm yields align with those that are publicly listed.

In Figure A.3, we report the relation between our imputed farm yields based on Equation (8) on the y-axis and those listed by PancakeSwap on the x-axis. All observations are nearly perfectly aligned with the (red dashed) 45-degree line. A linear projection of the imputed farm yields on the listed farm yields obtains a slope coefficient of 1.002 with an R^2 of 1.00. This is strong supportive evidence for the validity of our data building procedure.

5.2 Price, trade, and gas fee data

We obtain daily price (P_t^i) and trading volume ($V_{t,t+h}$) information for each farm. We source gas fee data from a proprietary data provider specialized in blockchain data services covering Bitcoin, Ethereum, Binance Smart Chain, among others. For a pair of cryptocurrency tokens in a liquidity pool, one typically serves as the numeraire, while the other is considered a token of interest. For example, in the ETH-BNB pool, ETH is the token of interest and BNB is the numeraire. In each liquidity pool, the price of the token of interest (e.g. ETH) is taken to be the most recent end-of-day price in GMT, where the price is expressed in terms of the numeraire token (e.g. BNB).

To find the prices of the numeraire tokens, we proceed in several steps. Table A.2 lists all numeraire tokens in the 219 liquidity pools and yield farms in our study. Among the

[6] https://pancakeswap.finance/farms

10 numeraire tokens, 4 are stablecoins pegged to the U.S. dollar: Binance USD (BUSD), TerraUSD (UST), Binance-Peg Tether (USDT), and Binance-Peg USD Coin (USDC). We collect their prices from CoinMarketCap for conversion. We infer the U.S. dollar value of the remaining numeraire tokens (e.g., Binance-Peg Ethereum (ETH)) from related token pairs of other liquidity pools in PancakeSwap. For example, we obtain the price of Binance-Peg Ethereum (ETH) in terms of Binance Coin (BNB) as numeraire, and the price of Binance Coin (BNB) in terms of Binance USD (BUSD) as numeraire. Using this approach, we can back out the price of Binance-Peg Ethereum (ETH) in Binance USD (BUSD) and convert it to U.S. dollars using the price of Binance USD in U.S. dollars from CoinMarketCap.

If there are no transactions on a given day, we use price information from the previous trading day instead. The daily trading volume of a pool is defined as the daily sum of trades across all cryptocurrencies in a given pool, measured in U.S. dollars.

We consider the impact of gas fees on the performance of yield farming strategies. Different functions executed by smart contracts incur different gas fees. To accurately impute the gas fees in the yield farming process, we first identify the chain of transactions that incur gas fees in a round-trip cost (see Table A.1). We then compute the average daily gas fee in U.S. dollars for each transaction in the chain. In a last step, we compute the gas fee for one round-trip cost by summing the average gas fee across all corresponding transactions.

5.3 Yield farmers data

We collect transaction data for each of the LP tokens through BscScan[7], a freely-accessible utility for searching data on the BSC. From these transaction logs, we then reconstruct the holding information of each wallet for each token. Transactions in which a user deposits cryptocurrency into a certain liquidity pool and receive LP tokens are represented as a LP token transfer from the null address to the wallet address of the user. Transactions in which a user stakes/unstakes their LP tokens in a yield farm are represented as a token transfer to/from the main staking contract. Transactions in which a user redeems their LP tokens at a liquidity pool in exchange for underlying tokens are represented in the data as a LP token transfer to the address of the LP token itself.

After collecting the above transaction data for all farms on PancakeSwap, we further refine the scope of our data in a few ways. First, we restrict our attention only to the accounts active in our sample period. Second, we eliminate wallet addresses belonging to non-PancakeSwap smart contracts, which may be a yield aggregator or an automated passive strategy. Finally, we omit wallet addresses that have transacted LP tokens with other, non-PancakeSwap smart contracts. These addresses employ staking on PancakeSwap as part of multi-platform investment strategies, which are beyond the scope of our current study. With the above refinements, we are left with 1,957,867 transactions made by 252,490 unique

[7] https://bscscan.com/

wallets active on PancakeSwap throughout our sample period. For accounts with a positive LP token balance at the end of our sample period, we make the standard assumption that these positions are exited on the last day of our sample. For each transaction, we match the price and offered yield of the LP token to the nearest end-of-day price by block height difference.

We compute several farmer-level measures of yield farming behavior. First, we compute *No. Farms*, the number of liquidity pools that each wallet interacts with. Second, we define *Efficiency* as *Time Staked*/*Time in Liquidity Pool* where *Time Staked* and *Time in Liquidity Pool* refer to the length of time for which the user has staked their LP tokens in a yield farm and the length of time for which the user has provided liquidity, respectively. We average this measure across liquidity pools. Third, we define *Staked Balance* and *LP Balance* as the time-weighted average balance for staking and liquidity provision. The prices used in these calculations are the nearest end-of-day price from the beginning of each holding period, and the weights are the length of each holding period. Finally, we define *Offered Farm Yield* of an yield farmer, as the time-weighted average of the offered yield at the beginning of each holding period. In order to guarantee that our results are not driven by very small investors, we choose farmers whose *LP Balance* is larger than $10. With this restriction, we are left with 207,699 unique farmers.

One caveat is that yield farmers may own and use multiple wallets. Hence, measures such as *No. Farms*, *Staked Balance*, and *LP Balance* could be underestimated. We believe that it is unlikely for yield farmers to systematically use multiple wallets for yield farming: there are no monetary benefits of doing this, and managing multiple wallets may add additional effort costs towards implementing yield farming strategies. We are also in the process of applying several wallet clustering algorithms to our sample in order to allay this concern.

5.4 Cryptocurrrency factors

Liu, Tsyvinski, and Wu (2019) document that a three-factor model using the cryptocurrency equivalents of the market, size and momentum factors are useful for explaining the cross section of expected cryptocurrency returns. We replicate these factors using their approach.

We obtain the cross-section of daily closing prices for cryptocurrencies from Coinmarketcap's historical API endpoint. For each cryptocurrency, prices are calculated using a volume-weighted average of prices reported from each of the markets for which Coinmarketcap has data. Our risk-free rate is from the St. Louis Fed's one-month constant maturity rate.

We exclude from our sample coins without trading volume data, coins with less than $1 million in market capitalization at the time of portfolio formation, and coins without price data for the following day. To control for potential outliers, we winsorize the market capitalization at the 1st and 99th percentiles during portfolio formulation.

For all three factors, we form portfolios at the end of the prior day and consider a one-day holding period before re-balancing at the end of the day. All returns are measured in terms of U.S. dollars. The excess cryptocurrency market return is constructed for a value-weighted portfolio of all coins with data on the portfolio formation day (prior to applying the filters), minus the risk-free rate at a daily frequency.

The excess cryptocurrency size factor is constructed as a zero-investment, long-short strategy with a long position defined as the value-weighted portfolio of coins in the bottom quintile of market capitalizations on the portfolio formation day, and a short position in the respective top quintile. For the cryptocurrency momentum factor, we further exclude coins for which the three-week price history is unavailable. The momentum factor is then constructed as a zero-investment, long-short strategy with a long position defined as the value-weighted portfolio of coins in the top quintile of three-week momentum on the portfolio formation day, and a short position in the respective bottom quintile.

In Appendix C.1, we describe our successful replication of Liu, Tsyvinski, and Wu (2019), suggesting that our cryptocurrency factors are reliably estimated.

6 Evidence

We first provide an overview of the data. We then describe the trading behavior of yield farmers and examine the risk and return characteristics of yield farming strategies. We end with a discussion on the interpretation of the evidence and the limitations of our analysis.

6.1 Descriptive overview

We illustrate in Panel (a) of Figure 5 the number of active farms during our sample period. We consider a farm to be active if the farm yield multiplier is nonzero, which means that liquidity providers staking LP tokens in the farm earn non-negative passive income.

We identify 219 unique active yield farms during our sample period, among which 110 farms co-exist both in both versions of PancakeSwap. The remaining 109 farms are active in only one of both versions. The total number of active farms at any period increases quickly from inception of PancakeSwap to a peak of 160 farms in July 2021.

In Panel (b) of Figure 5, we plot the the Total Value Locked (TVL) in all active farms, i.e., the aggregate amount of liquidity deposited for yield farming. Yield farming at PancakeSwap has experienced extraordinary growth, with TVL surpassing \$7 billion in May 2021. Analogously to the boom and bust cycles experienced by Bitcoin and other cryptocurrency markets, TVL dropped sharply following its peak and experienced renewed momentum after the drop in aggregate liquidity.

In Figure 6, we provide a histogram of yield farm duration. We define the duration of a yield farm to be the period during which the farm offers a non-negative passive income to yield farmers, that is, the farm's multiplier is strictly positive. We find a significant amount of heterogeneity for the life span of yield farms. The mean (median) duration is 130.22 (111) days with a standard deviation of 84.88 days. Some farms have very short durations.

Each yield farm features a unique pair of cryptocurrency tokens. We report in Table 3 the ten largest yield farms in terms of their TVL as of September 5, 2021. The largest farm is associated with CAKE-BNB and recorded TVL of $931.72 million in its pool. The 10th largest farm is associated with USDC-USDT, which started on June 28, 2021, and recorded TVL of $103.29 million by the end of our sample period.

There is a significant amount of heterogeneity in the yields that are offered by the ten largest farms reported in Table 3, ranging from 5.82% for USDC-BUSD to 34.90% for CAKE-BNB. These headline yields look attractive on an annualized basis.

In Panel (a) of Figure 2, we plot the time-variation in the median farm yield together with its cross-sectional distribution. In Panel (b), we report the same evidence from November 1, 2020 onwards, due to extreme yields offered during the initial phases of yield farming. These figures highlight significant fluctuations in the median farm yield, which is often higher than 100%. In addition, there is significant variation in dispersion of farm yields, as is underscored by the fluctuations in the interquartile range of the yield farm distribution. Such rich variation in yields across farms and across time provides an opportunity to better understand the drivers of cross sectional variation in the risk and return characteristics of yield farming strategies and the performance of liquidity provision.

In Table 4, we report the summary statistics of the return performance associated with yield farming strategies. In Panel A, we focus on annualized returns computed for a daily trading horizon. The average (median) return to yield farming is 147.7% (193.16%) during our sample period. This is the average return across the 219 unique yield farms, which have a duration that is about 129 days, on average. Returns to yield farming are volatile with a standard deviation that is on average 125.89%. Yield farming generates a return performance that is negatively skewed (-0.3061), fat-tailed (7.6484) and weakly negatively serially correlated with a first-order autocorrelation coefficient of -0.0999.

In Panel B of Table 4, we provide the same statistics for a weekly trading horizon. the average yield farm has a duration of 20 weeks. The key difference at the weekly frequency is that yield farming performance is weakly positively serially correlated (AC1 coefficient of 0.0407), less negatively serially correlated and less fat tailed.

We also provide information about the returns to liquidity provision that excludes the staking of liquidity tokens into yield farms (liquidity mining). The returns to liquidity mining alone are significantly smaller with average (median) annualized returns of 19.44% and 6.30% (57.53% and 53.48%) at the daily and weekly frequency, respectively.

Another useful comparison is the return performance of a simple buy-and-hold strategy that invests into the pair of cryptocurrency tokens associated with a pool. This comparison is useful because investors face a choice of directly investing into a pair of cryptocurrency tokens or stake them to a liquidity pool. At the daily (weekly) frequency, a buy-and-hold strategy earned on average 49.10% (37.59%) on an annualized basis during our sample period. Thus, buy-and-hold strategies earn on average about a third of the return performance of a yield farming strategy, before considering the costs associated with each strategy.

The returns to yield farming may seem extraordinarily large at first. This is driven by our sample period that overlaps with a strong bull market in asset markets in general, and for cryptocurrencies in particular. We illustrate this by comparing the return performance of yield farming strategies to other well-known strategies and list that information under "benchmark strategies" in Table 4.

During the same period, Bitcoin and Ethereum earned an average of 168.72% and 266.41% on an annualized basis for a daily trading horizon. In comparison, the cryptocurrency market factor of Liu, Tsyvinski, and Wu (2019) earned 186.27% at the daily frequency, and the S&P 500 index recorded a performance of 31.70%.[8] We also find that the average return for MVIS CryptoCompare Digital Assets 10 (100) index is 227.95% (219.20%) on an annualized basis for a daily trading horizon.

6.2 Evidence of lack of investor sophistication

6.2.1 Evidence from yield farmers data

Several features of the yield farming infrastructure at PancakeSwap enable us to infer information about the activities of its participants. In this subsection, we provide some evidence consistent with cross-sectional variation in sophistication across yield farmers.

First, PancakeSwap migrated to a new version on April 24, 2021 when it upgraded the technological and security features of its smart contract design. Since then, liquidity pools and yield farms associated with a particular pair of cryptocurrency tokens have coexisted on both old and new platforms. Liquidity providers were strongly encouraged to switch their liquidity provision from version 1 to version 2, but had to trigger the switch themselves.

In Figure 7, we show that a significant amount of liquidity remains in the liquidity pools associated with the old version. This is puzzling since the switch to the new version is considered to be a strictly dominant investor strategy. Migrating liquidity to the new version delivers higher rewards for staking the same tokens as in version 1, alongside lower transaction costs. This is likely a sign of investor inertia or inattention.

[8] We present detailed statistics for the construction of all cryptocurrency factors of Liu, Tsyvinski, and Wu (2019) in Tables A.3 and A.4.

Second, returns to yield farming involve several independent transactions. Investors first need to provide liquidity to liquidity pools. The liquidity tokens that certify the liquidity provision then need to be independently staked to a yield farm. Combining both transactions is a strictly dominant strategy compared to liquidity provision alone. However, we show in Figure 8 that the number of LP tokens staked in yield farms is significantly lower than the aggregate amount of LP tokens minted to certify liquidity provision.

We would expect the staking ratio to be equal to one at all times. However, the median ratio is below one most of the time. The 10th (25th) percentile of the distribution even drops to as low as 30% (85%). This is further evidence that supports the lack of investor sophistication in this market. However, we caveat this interpretation with the possibility of investors staking their LP tokens in third-party yield farm aggregators. While we currently do not have access to this information, we are in the process of collecting it.

In order to better understand this phenomenon of forgoing farming opportunities, we look to the trading data for individual yield farmers. Table 7 Panel A shows farmer-level summary statistics. The average yield farmer invests in 1.81 farms and provides $6,959 of liquidity. However, *Staked Balance*, the dollar value of LP tokens staked in yield farms, is substantially lower than *LP Balance*, the provided liquidity. This suggests that a significant profit generated from farming is lost for investors who miss the farming opportunities possibly due to the complex nature of trading strategies. Consistent with this finding, *Efficiency* is 0.82, which implies that the length of time that the average yield farmer keeps his/her LP tokens in a farm is significantly lower than the length of time that he/she keeps corresponding liquidity in a liquidity pool.

In Panel B of Table 7, we decompose the farmers into two groups depending on *Efficiency* using 0.98 as a threshold. Interestingly, we find that a large number of farmers do not meet this efficiency requirement. 60,514 farmers' *Efficiency* measures are less than or equal to 0.98. Although the average *LP Balance* of the low efficiency group is lower than that of the high efficiency group, we find that average *LP Balance* of the low efficiency group remains economically significant ($3,099). Their average efficiency is also remarkably low (0.40), and their *Offered Farm Yield* is relatively high (140%). Overall, the results suggest that a large number of yield farmers miss a substantial amount of revenue from farming.

In Panel C, we decompose the farmers into quintiles based on their average *LP Balance*. We observe that *LP Balance* is positively correlated with *Efficiency*, suggesting that smaller yield farmers are more likely to miss their farm yields, which range from 131% to 153% across quintiles. Nevertheless, we still observe a significant fraction of yield farmers missing farm yields, even in the top quintile. Another interesting observation is that differences in average *LP Balance* are substantial across quintiles. For instance, the average *LP Balance* of quintile 1 is only $25.55, whereas that of quintile 5 is $33709. This suggests that there exists significant cross-sectional dispersion in average investment sizes among users of PancakeSwap. Both very large and small investment sizes could be concerning because they may generate sub-optimal performance, as illustrated in Section 6.4.

6.2.2 Additional suggestive evidence

There exists additional evidence suggesting that a significant fraction of yield farmers are not sophisticated. First, there is suggestive evidence that many investors in PancakeSwap are small retail investors. According to DappRadar[9], PancakeSwap registered 435,130 active users on October 24, 2021, in contrast to 47,730 active users recorded for Uniswap. The number of active users on PancakeSwap is the largest among all decentralized applications built on all blockchains tracked by DappRadar. The trading volume in PancakeSwap was about $1.2B on October 24, which implies that the average yield farmer in PancakeSwap traded $2,757.

In addition, survey results suggest that a significant fraction of yield farmers are not sophisticated. CoinGecko, a major cryptocurrency data provider, surveyed 1,347 cryptocurrency investors regarding yield farming in August 2020. (CoinGecko, 2020) Interestingly, a significant fraction of yield farmers seem to be overconfident and unsophisticated. According to the survey, 79% of yield farmers claim to understand the associated risks and rewards of yield farming to a reasonable extent. However, about 40% of yield farmers report that they could not read smart contracts to verify potential vulnerabilties or scams of the yield farms. In addition, 33% of yield farmers do not know what the impermanent loss is, implying that they are taking risks that they are unaware of.

6.3 Performance of yield farming strategies without transaction costs

In Table 5, we decompose yield farming returns into their four components: capital gains, impermanent losses, trading fees, and farm yields. We first focus on the full sample results.

Farm yields contribute the most to yield farming performance, with an average daily log return of 153.92%. Capital gains are the second largest contributor, with an annualized daily log return that is 38.29% on average. This is comparable to the return performance of the S&P500 index during that period.

We note that capital gains are significantly more volatile than farm yields, and that they have more extreme negative and positive outcomes. The annualized standard deviation is 124.73%, compared with 3.08% for farm yields, and the wider interquartile range for capital gains reflects the greater kurtosis of 7.72, compared to a a distribution that is much less fat-tailed for farm yields. The persistence of returns is also different across these two components. While capital gains exhibit weak negative serial correlations, farm yields are persistent with a first order autocorrelation coefficient of 0.7625.

The annualized daily impermanent loss is -41.08% on average. In addition, the distribution is negatively skewed and exhibits the largest excess kurtosis among all four components, a

[9]DappRadar: https://dappradar.com/rankings

value of 49.48. This reflects investors' negative exposure to correlation risk, since impermanent losses are exponentially sensitive to the return divergence between the underlying pairs of cryptocurrency tokens.

The annualized daily trading fee is 11.36%, on average, making it the least important contributor to yield farming performance. Despite the lower volatility (standard deviation of 0.72%), trading fees can become important, as demonstrated by the positive skewness (2.9542) and kurtosis (18.4164).

In Table 5, we report similar statistics for yield farms, sorted into terciles by the magnitude of their average in-sample offered yield. This sorting exercise reveals a negative relationship between the headline yields and capital gains performance.

In farms with low headline yields (Tercile 1), capital gains are significantly larger than farm yields (152.12% vs. 51.20%). On the other hand, in farms with high headline yields, capital gains are highly negative, on average (-84.18%), while farm yields deliver an annualized daily return of 271.96%. Trading fees and impermanent losses appear similar across all three terciles.

In Panel B of Table 5, we report the decomposed annualized return performance for weekly trading horizons. The patterns are broadly similar to those of daily trading horizons.

In Table 6, we perform a variance decomposition. This helps better understand the proportion of variation in the aggregate return series explained by each of the four return components arising from capital gains, impermanent losses, trading fees, and farm yields. For the variance decomposition, we split the covariance terms equally for each component. We report the variance decomposition for the overall population, and for each tercile of the population sorted by their average offered yield throughout the sample period.

Consistent with the large standard deviations observed for capital gains in Table 5, capital gains dominate the overall variation of the raw yield farming performance. The proportion of variation explained by capital gains is close to 100%. This is true at both the daily and weekly trading horizons.

We assess the value-weighted performance of yield farming strategies in Table 8, using the pools' aggregate liquidity as weighting factors. We take the perspective of a U.S. investor who starts from an initial hypothetical $1 USD investment and ignore all transaction costs. We compute returns in excess of the three-month U.S. Treasury bill secondary market rate. We focus on the daily trading frequency in Panel A.

We find that, without transaction costs, yield farming was highly profitable during our sample period. The average value-weighted excess return delivered an annualized return of 237.12%. This is significantly larger than the returns to a strategy that focuses only on liquidity mining (185.98%), and larger than a buy-and-hold strategy in the same pairs of cryptocurrency tokens associated with the liquidity pools (196.28%). All three strategies deliver negatively skewed performances, with a non-trivial amount of excess kurtosis.

To assess risk-return trade-offs, we standardize the return performance by the annualized standard deviations and compute Sharpe ratios for all investment strategies. These measures suggest a lucrative risk-return trade-off, with values ranging from 2.38 for buy-and-hold strategies to 2.89 for yield farming.

While Sharpe ratios may appear extraordinarily large, they are comparable to similar magnitudes recorded during our sample period for benchmark trading strategies. For example, an investment in the S&P 500 Index delivered a Sharpe ratio of 2.37 during the same period. Sharpe ratios for other comparable investments range from 1.97 for the cryptocurrency market factor of Liu, Tsyvinski, and Wu (2019) to 2.41 for an investment in Ethereum. It is important to note that, for reasons of simplicity and clarity, we do not account for autocorrelation in our annualization of return volatility. At a weekly measurement frequency, for instance, yield farming strategies have large and positive autocorrelation coefficients. Correcting for them will increase the annualized standard deviation, and thereby decrease our reported Sharpe ratios for yield farming strategies. Since these coefficients are much larger for yield-farming strategies at a weekly frequency compared to our benchmarks, the overall effect of the correction will worsen the relative performance of yield farming strategies.

We also report alphas estimated using the three-factor cryptocurrency return model of Liu, Tsyvinski, and Wu (2019). Their framework suggests that a three-factor model with cryptocurrency market, size, and momentum factors can price the cross-section of cryptocurrency returns. Thus, we assess the risk-adjusted performance of yield farming performance relative to this three-factor cryptocurrency benchmark. We find that the alpha for yield farming investments is on average 137.78%. Because of the short and volatile sample period, this alpha is estimated with a t-statistic of only 1.86.

6.4 Performance of yield farming strategies with transaction costs

The evidence suggests that yield farming delivers attractive returns, with high risk adjusted returns and Sharpe ratios. We question whether these returns are realistically attainable, despite the positive bull run observed during our sample period. An important insight of our study is the careful examination of trading fees and the price impact implicit in staking cryptocurrency pairs to liquidity pools and in harvesting farm yields.

In Table 9, we document the performance of yield farming strategies, accounting for gas fees, trading fees and price impact. We then compare these results to the frictionless benchmark. We add the information ratio using the best fit from the three-factor model of Liu, Tsyvinski, and Wu (2019) as a benchmark portfolio to compute tracking errors.

Despite the significantly lower transaction costs recorded on BSC compared to Ethereum, gas fees significantly lower the return performance. This is because the multiplicity of transactions that are needed for a round-trip transaction can accumulate to non-trivial amounts, especially with frequent rebalancing.

In Panel A of Table 9, we show the performance of yield farming strategies after accounting for transaction cots. We consider initial investments ranging from \$5,000 to \$1,000,000.[10]

The impact of gas fees and trading costs is especially harmful for small size investments, since they are based on flat dollar amounts. When the investment size is too small, the fixed transaction costs reflect a large proportion of the investment so that they absorb a large fraction of the positive return performance. This incentivizes larger investment amounts to reduce the dollar cost basis. However, amounts of \$100,000 or more may not be an option for unsophisticated retail investors and we find that a large proportion of investors invest less than \$1,000 in farms in Section 6.2.

On the other hand, when the investment size is too large, there is too much capital relative to the liquidity provision ability of a pool. Thus, when swapping to get the tokens, the slippage from illiquidity is too high. We previously discussed that larger investments endogenously also lead to lower farm yields, thereby putting further downward pressure on the investment performance. Across the board, we notice that the risk-adjusted performance becomes negative, as suggested by the negative alphas and information ratios, regardless of the investment size.

These observations also have implications for diversification. A portfolio of fewer yield farms would save more on fixed transaction costs, but would be more exposed to illiquidity (slippage) when opening/closing positions, due to higher idiosyncratic risk. In contrast, holding a more diversified portfolio of farms would cost more but would lower potential losses from illiquidity (slippage) when opening/closing positions.

To shed more light on the precise mechanism of these trade-offs, we plot in Figure 10 the hypothetical return performance of two diversified yield-farming investment strategies as a function of investment size. We consider both a value-weighted and an equal-weighted strategy, whereby the value-weighted (equal-weighted) portfolio consists of all yield farm returns, beginning on October 20, 2020 and re-balancing every seven days thereafter until September 5, 2021.

In Panels (a.1), (a.2) and (a.3), we observe a non-monotonic relation between investment size and return performance in PancakeSwap. For small investment sizes, returns are negative and volatile, leading to negative Sharpe ratios. The same phenomenon appears for larger investments sizes close to \$10 million. Maximum Sharpe ratios are attained for investment amounts ranging between \$100,000 and \$1 million for value-weighted and equal-weighted portfolios, respectively.

[10]Note that the average return may be less than -100% because the average return is computed based on log rather than arithmetic returns.

Our analysis in Table 9 shows that, besides the drop in yield farming performance, the impact on performance is dependent on the farm yield distribution. In particular, we observe the greatest reduction in yield farm performance for farms that offer the highest headline yields (tercile 3). That impact is especially pronounced for large investments such as $100,000 and $1,000,000. This important observation leads us to further assess the relation between flow and performance, since there is important evidence from other asset markets that suggest investors reach for yield (e.g., Becker and Ivashina, 2015; Choi and Kronlund, 2018; Chen and Choi, 2021; Bordalo, Gennaioli, and Shleifer, 2016) and consequentially pursue investment strategies with large headline rates (e.g., Henderson and Pearson, 2011; Célérier and Vallée, 2017; Egan, 2019; Henderson, Pearson, and Wang, 2020; Shin, 2021).

In Panel B of Table 9, we report the yield farming performance when we account for trading frictions at a weekly trading horizon. The impact on performance is naturally less dramatic, since portfolio re-balancing happens seven times less frequently. However, we observe the same pattern in that the biggest price impact is suffered by the farms with the highest headline rates and for large investments. Most importantly, we observe that Sharpe ratios are in general lower than those of benchmark indexes such as the S&P 500 Index and MVIS 10 Index.

Since we study yield farming at PancakeSwap, we believe our results to be conservative. We show in Figure 9 the average gas fee incurred in yield farming at PancakeSwap and at SushiSwap, one of the largest DEX yield farms in Ethereum as of September 5, 2021. For PancakeSwap, the average cost to enter (exit) yield farming over all days is $1.39 ($1.89). For SushiSwap, the average cost to enter (exit) over all days is $109.12 ($164.34). Thus, any impact from transaction costs will likely be amplified for alternative studies based on SushiSwap.

We go through the similar data collection process to construct the performance of SushiSwap-based yield farming strategies. Consistent with our intuition, in Panel (b) of Figure 10, we find that the performance of yield farming strategies in SushiSwap is poor. Both maximum average return and maximum Sharpe ratio of SushiSwap-based yield farming strategies are significantly lower than those of PancakeSwap-based strategies. For example, the maximum Sharpe ratio is lower than 1.5 and the Sharpe ratio is particularly low when the size of investment is low due to high gas cost. The maximum Sharpe ratio is attained at the investment size of $4-5 million and even at this size of investment, the Sharpe ratio is lower than that of S&P 500 (2.4).

6.5 The relation between farm flows and performance

In light of the observation that price impact is largest for farms with high headline yields, we assess the relation between farm flows and performance. To compute farm flows, we closely follow the mutual fund literature and estimate them using the time series variation in each pool's aggregate liquidity L_t and the per period farm growth due to return performance $R_{t,t+h}$ (e.g., Sirri and Tufano, 1998; Coval and Stafford, 2007). We provide a detailed description in Section 4. We aggregate flows at the daily frequency to obtain weekly flows. In Table 10, we report the results from a regression of farm flows on offered farm yield (y_t^j), lagged farm flows, and past performance of a yield farming strategy. Specifically, we regress

$$Flow_{t,t+7}^j = a + by_t^j + \sum_{k=1}^{K} c_k \cdot Flow_{t-7k,t-7(k-1)}^j + \sum_{h=1}^{H} d_k R_{t-7h,t-7(h-1)}^j + e^\top X_t + f_j + \varepsilon_t^j, \quad (13)$$

including pool fixed-effects and common crypto market factors X_t as control variables.

The first notable observation is that yield farmers seem to reach for high yields. The coefficients for *Offered Farm Yield* are about 0.0149 in columns (1) and (3), both of which are statistically significant at the 5% level. The magnitude is economically significant as well. A one-standard-deviation increase in *Offered Farm Yield* is associated with a 0.012 (1.2%) increase in one-week ahead *Flow*. The effect becomes even more significant, both statistically and economically, once we add farm fixed effects to control for farm-specific characteristics, such as underlying tokens in the pool. The coefficient is 0.0973 (0.114) in column (4) ((6)) and it is statistically significant at the 1% level. Under these specifications, a one-standard-deviation increase in *Offered Farm Yield* is associated with 0.077-0.091 (7.7% - 9.1%) increase in *Flow*. This is an economically significant magnitude.

High yield seeking behavior is observed in many other financial markets (Henderson and Pearson, 2011; Becker and Ivashina, 2015; Bordalo, Gennaioli, and Shleifer, 2016; Célérier and Vallée, 2017; Choi and Kronlund, 2018). A vast majority of this line of research emphasizes the role of intermediaries as the source of the reaching-for-yield phenomenon. In contrast, yield farming operates through smart contracts on decentralized markets without financial intermediation. Our evidence suggests that reaching for yield may also exist in decentralized finance. This is of interest since our results imply that reaching for yield can arise even in the absence of financial intermediaries and related agency conflicts or competing incentives.

Another important finding is that yield farmers chase past performances of yield farming strategies. The one-week ahead *Log return* is a strong predictor of next week's *Flow*, similar to what is observed for mutual funds (e.g. Brown, Harlow, and Starks, 1996; Chevalier and Ellison, 1997; Berk and Green, 2004; Sirri and Tufano, 1998). This is similarly intriguing in the absence of financial intermediation and agency frictions, since financial intermediaries are often cited as important drivers of the positive relation between flows and past returns (e.g. Chevalier and Ellison (1997); Berk and Green (2004)).

Finally, Figure 11 highlights a notable relationship between past *Flow* and future performance of yield farming strategies. Although superior past returns and high offered yields can generate high inflows of liquidity, high inflows do not lead to better performance of yield farming strategies. In fact, higher inflows predict lower returns from yield farming strategies in the next 2-3 weeks. This explains a lack of persistence in yield farming returns.[11] Berk and Green (2004) propose a model with rational learning and decreasing returns-to-scale to justify the lack of persistence. In this model, to generate decreasing return-to-scale, it is crucial to assume that the cost of managing a fund of a certain size is increasing, convex function with respect to the size of the fund. In yield farming, while higher *Flow* does not increase any managial costs, it mechanically drives down offered farm yields, due to the logic implemented in a smart contract, which can eventually lead to under-performance in the long run.

7 Conclusion

We provide the first characterization of yield farming, a novel decentralized financial service available to retail investors in the cryptocurrency ecosystem. Using a novel hand-collected data set on 219 yield farms from PancakeSwap, an automated market maker operating on the Binance Smart Chain, we also assess the yield farming return performance and describe the associated risks.

While yield farming appears to deliver positive investment performances during our sample period that is comparable to other standard investment strategies, Sharpe ratios are significantly reduced after accounting for transaction fees and price impact. With daily rebalancing, risk-adjusted returns become negative. Investors are also exposed to large losses that are driven by the return differential of underlying cryptocurrency pairs associated with yield farms.

We uncover a non-monotonic trade-off between investment size and return performance. Small trades are penalized by high nominal transaction costs and low liquidity is associated with volatile returns. Large trades are less penalized by excessive gas fees, but too much liquidity provision may lead to price impact and slippage that hurts investors and amplifies volatility. We find that maximum and non-negative Sharpe ratios arise for investment sizes between $100,000 and $1 million. Thus positive yield farming performance may not be attainable to small retail investments.

Importantly, we find that the price impact is larger for farms that offer higher yields. Our analysis suggests that this is because flows into farms chase past positive performance. However, this leads to greater price impact that ultimately lowers excess returns.

Our analysis helps understand the risk and return characteristics of yield farming. Since this is a complex investment strategy that is easily accessible to retail investors and subject to significant downside risks, we believe that our findings are helpful to regulators in determining the need for better risk disclosure and investor protection.

[11] In an untabulated study, we regress the one-week ahead yield farming return on past one-week return with farm fixed effects and find a regression coefficient of 0.0084 that is statistically insignificant, which implies that yield farming performance is not persistent.

References

Aigner, Andreas A., and Gurvinder Dhaliwal, 2021, UNISWAP: Impermanent Loss and RiskProfile of a Liquidity Provider, *SSRN Working Paper 3872531.*

Angeris, Guillermo, Hsien-Tang Kao, Rei Chiang, Charlie Noyes, and Tarun Chitra, 2019, An Analysis of Uniswap Markets, *Working paper.*

Aoyagi, Jun, 2021, Liquidity Provision by Automated Market Makers, *Working paper.*

Aoyagi, Jun, and Yuki Ito, 2021, Coexisting Exchange Platforms: Limit Order Books and Automated Market Makers, *working paper.*

Augustin, Patrick, Alexey Rubtsov, and Donghwa Shin, 2021, The Impact of Derivatives on Cash Markets: Evidence from the Introduction of Bitcoin Futures Contracts, *Working paper.*

Becker, Bo, and Victoria Ivashina, 2015, Reaching for Yield in the Bond Market, *Journal of Finance* 70, 1863–1902.

Berk, Jonathan B, and Richard C Green, 2004, Mutual Fund Flows and Performance in Rational Markets, *Journal of Political Economy* 112, 1269–1295.

Biais, Bruno, Christophe Bisiere, Matthieu Bouvard, and Catherine Casamatta, 2019, The Blockchain Folk Theorem, *Review of Financial Studies* 32, 1662–1715.

Biais, Bruno, Christophe Bisiere, Matthieu Bouvard, Catherine Casamatta, and Albert J Menkveld, 2020, Equilibrium bitcoin pricing, *Working paper.*

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2016, Competition for Attention, *Review of Economic Studies* 83, 481–513.

Borri, Nicola, and Kirill Shaknov, 2021, The Cross-Section of Cryptocurrency Returns, .

Brown, Keith C, W Van Harlow, and Laura T Starks, 1996, Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry, *Journal of Finance* 51, 85–110.

Capponi, Agostino, and Ruizhe Jia, 2021, The Adoption of Blockchain-based Decentralized Exchanges, *Working paper.*

Célérier, Claire, and Boris Vallée, 2017, Catering to Investors through Security Design: Headline Rate and Complexity, *Quarterly Journal of Economics* 132, 1469–1508.

Chen, Qianwen, and Jaewon Choi, 2021, Reaching for Yield and Bond Returns, *Working paper.*

Chevalier, Judith, and Glenn Ellison, 1997, Risk Taking by Mutual Funds as a Response to Incentives, *Journal of Political Economy* 105, 1167–1200.

Choi, Jaewon, and Mathias Kronlund, 2018, Reaching for Yield in Corporate Bond Mutual Funds, *Review of Financial Studies* 31, 1930–1965.

Choi, Kyoung Jin, Alfred Lehar, and Ryan Stauffer, 2018, Bitcoin Microstructure and the Kimchi premium, *Working Paper*.

CoinGecko, 2020, Yield Farming Survey 2020, *CoinGecko.com*.

Cong, Lin William, and Zhiguo He, 2019, Blockchain Disruption and Smart Contracts, *The Review of Financial Studies* 32, 1754–1797.

Cong, Lin William, Xi Li, Ke Tang, and Yang Yang, 2021, Crypto Wash Trading, *Working paper*.

Cong, Lin William, Ye Li, and Neng Wang, 2021, Tokenomics: Dynamic adoption and valuation, *Review of Financial Studies* 34, 1105–1155.

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of FInancial Economics* 86, 479–512.

Dyhrberg, Anna Haubo, 2020, One Coin, Many Markets - How Market Frictions A ect Arbitrageurs, .

Easley, David, Maureen O'Hara, and Soumya Basu, 2019, From Mining to Markets: The Evolution of Bitcoin Transaction Fees, *Journal of Financial Economics* 134, 91–109.

Egan, Mark, 2019, Brokers versus Retail Investors: Conflicting Interests and Dominated Products, *Journal of Finance* 74, 1217–1260.

Foley, Sean, Jonathan R Karlsen, and Talis J Putnins, 2019, Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed through Cryptocurrencies?, *The Review of Financial Studies* 32, 1798–1853.

Gandal, Neil, JT Hamrick, Tyler Moore, and Tali Oberman, 2018, Price Manipulation in the Bitcoin Ecosystem, *Journal of Monetary Economics* 95, 86–96.

Gensler, Gary, 2021, Remarks Before the Aspen Security Forum, Remarks by Chairman Gary Gensler at the Annual Aspen Security Forum.

Ghent, Andra C., Walter N. Torous, and Rossen I. Valkanov, 2019, Complexity in Structured Finance, *Review of Economic Studies* 86, 694–722.

Griffin, John M, and Amin Shams, 2020, Is Bitcoin Really Untethered?, *Journal of Finance* 75, 1913–1964.

Han, Jianlei, Shiyang Huang, and Zhuo Zhong, 2021, Trust in DeFi: An Empirical Study of the Decentralized Exchange, *Working paper*.

Harvey, Campbell R, 2016, Cryptofinance, *Working paper.*

Harvey, Campbell R, Ashwin Ramachandran, and Joey Santoro, 2021, DeFi and the Future of Finance, *Working paper.*

Hautsch, Nikolaus, Christoph Scheuch, and Stefan Voigt, 2019, Limits to Arbitrage in Markets with Stochastic Settlement Latency, *SSRN Working Paper 3302159.*

Henderson, Brian J, and Neil D Pearson, 2011, The Dark Side of Financial Innovation: A Case Study of the Pricing of a Retail Financial Product, *Journal of Financial Economics* 100, 227–247.

Henderson, Brian J, Neil D Pearson, and Li Wang, 2020, Pre-trade hedging: Evidence from the Issuance of Retail Structured Products, *Journal of Financial Economics* 137, 108–128.

Hinzen, Franz J, Kose John, and Fahad Saleh, 2020, Bitcoin's Fatal Flaw: The Limited Adoption Problem, *NYU Stern School of Business.*

Howell, Sabrina T, Marina Niessner, and David Yermack, 2020, Initial Coin offerings: Financing Growth with Cryptocurrency Token Sales, *Review of Financial Studies* 33, 3925–3974.

Hu, Albert S, Christine A Parlour, and Uday Rajan, 2019, Cryptocurrencies: Stylized Facts on a New Investible Instrument, *Financial Management* 48, 1049–1068.

Kroeger, Alexander, and Asani Sarkar, 2017, The Law of One Bitcoin Price?, *New York Fed Working Paper.*

Krückeberg, Sinan, and Peter Scholz, 2020, Decentralized Efficiency? Arbitrage in Bitcoin Markets, *Financial Analysts Journal* 76, 135–152.

Lee, Jongsub, Tao Li, and Donghwa Shin, 2022, The Wisdom of Crowds in FinTech: Evidence from Initial Coin Offerings, *The Review of Corporate Finance Studies* 11, 1–46.

Lehar, Alfred, and Christine A Parlour, 2021, Decentralized Exchanges, *Working paper.*

Li, Tao, Donghwa Shin, and Baolian Wang, 2021, Cryptocurrency Pump-and-Dump Schemes, *Working paper.*

Li, Ye, and Simon Mayer, 2021, Money Creation in Decentralized Finance: A Dynamic Model of Stablecoin and Crypto Shadow Banking, .

Liu, Yukun, Aleh Tsyvinski, and Xi Wu, 2019, Common Risk Factors in Cryptocurrency, *Forthcoming in Journal of Finance.*

Makarov, Igor, and Antoinette Schoar, 2019, Price Discovery in Crypto Currency Markets, *AEA Papers & Proceedings.*

Makarov, Igor, and Antoinette Schoar, 2020, Trading and Arbitrage in Cryptocurrency Markets, *Journal of Financial Economics* 135, 293–319.

Neuder, Michael, Rithvik Rao, Daniel J. Moroz, and David C. Parkes, 2021, Strategic Liquidity Provision in Uniswap v3, *working paper.*

Odean, Terrance, 1999, Do Investors Trade Too Much?, *American Economic Review* 89, 1279–1298.

Oliver, Joshua, 2021, Traders lend out cryptocurrencies in quest for huge returns, .

Pagnotta, Emiliano, and Andrea Buraschi, 2018, An Equilibrium Valuation of Bitcoin and Decentralized Network Assets, *Working paper.*

Park, Andreas, 2021, The Conceptual Flaws of Constant Product Automated Market Making, *Working paper.*

Saleh, Fahad, 2021, Blockchain Without Waste: Proof-of-Stake, *Review of Financial Studies* 34, 1156–1190.

Shin, Donghwa, 2021, Extrapolation and Complexity, *Working paper, University of North Carolina.*

Sirri, Erik R, and Peter Tufano, 1998, Costly Search and Mutual Fund Flows, *Journal of Finance* 53, 1589–1622.

Sockin, Michael, and Wei Xiong, 2020, A Model of Cryptocurrencies, *Working paper.*

Yermack, David, 2017, Corporate Governance and Blockchains, *Review of Finance* 21, 7–31.

Yu, Yang Gloria, and Jinyuan Zhang, 2018, A Revisit to Capital Controls Policies When Bitcoin Is in, *SSRN Working Paper 3053474.*

Figure 1: Growing Popularity of Decentralized Finance

In this figure, we plot the total value locked (TVL) of all decentralized finance platforms in billions of dollars, as reported by DeFiLlama. The figure starts on January 1, 2020 and ends on September 5, 2021.

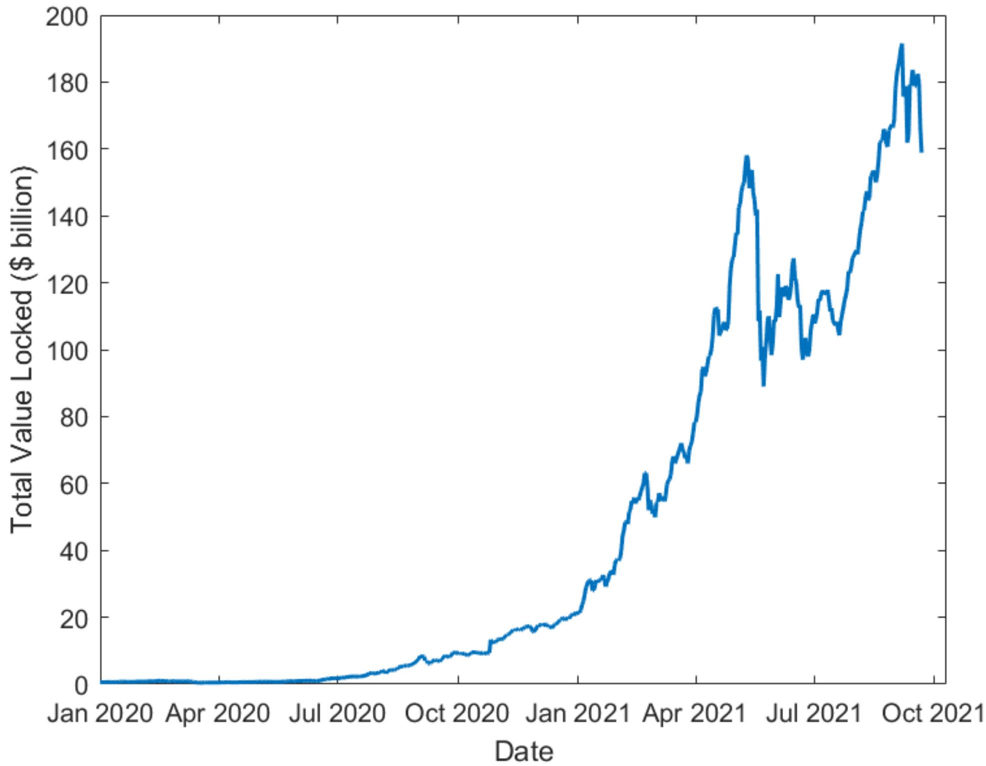
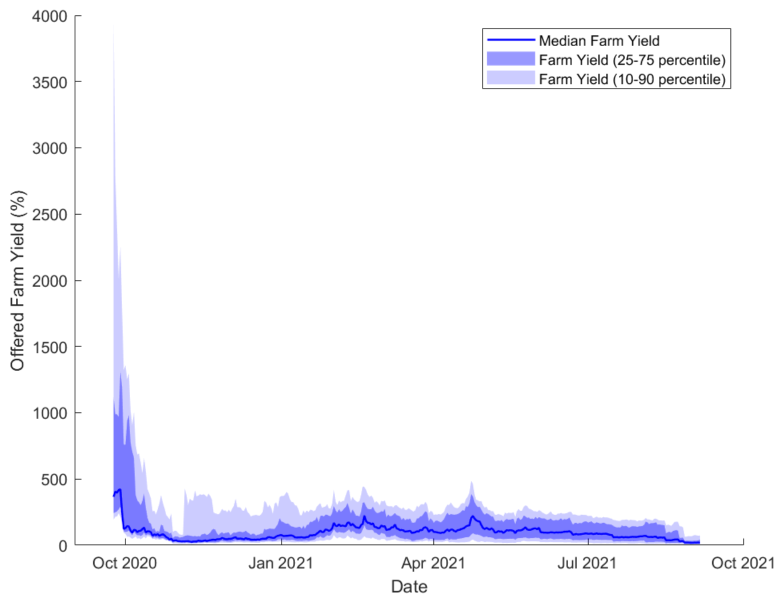


Figure 2: Offered Farm Yields

In this figure, we plot the annualized farm yields offered to yield farmers. In Panel (a), we provide the historical annualized offered farm yields during the period between September 23, 2020 and September 5, 2021. In Panel (b), we restrict our sample period to extend from October 20, 2020 to September 5, 2021. The solid blue line indicates the median annualized offered farm yield. Dark and light shaded areas represent the interquartile range, as well as the 10th and 90th percentiles of the yield farm distribution, respectively.

(a)



(b)

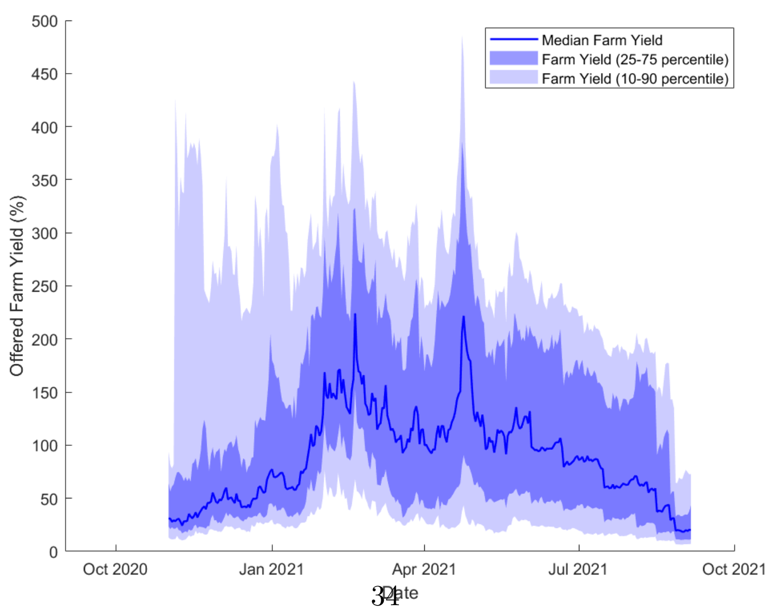
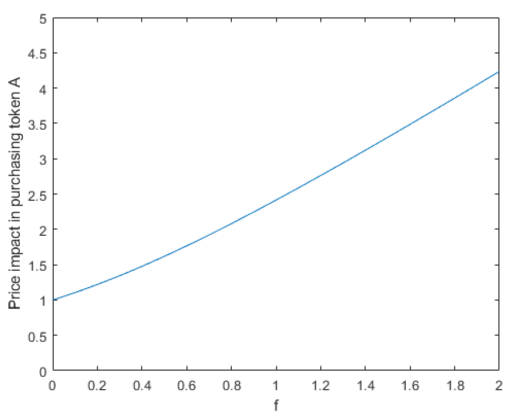


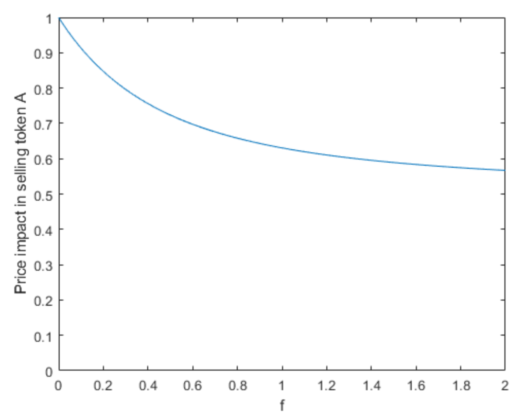
Figure 3: Model-Implied Price Impact due to Yield Farming

In this figure, we illustrate how the size of investment in yield farming creates price impact, which affects returns from yield farming. The parameter f defines the relative ratio of the size of the investment to the size of the liquidity pool, i.e. investment/size of liquidity pool (I_t/L_t). Consider two cryptocurrencies A and B in a liquidity pool with token B being the numeraire token such as BNB or BUSD. Panel (a) shows the relation between f and the price impact on token A when purchasing token A for providing liquidity (together with token B) to a pool. The y-axis plots the multiple to the current price of token A in U.S. dollars. A value of 2 implies that a yield farmer would have to pay twice the current market price of token A to acquire it for liquidity provision. Panel (b) plots the relation between f and the price impact on token A when selling it after liquidity withdrawal from the pool. Panel (c) plots the impact of investment size on gross returns from capital gain and impermanent loss. For example, $\lambda(f) = 0.5$ implies that the gross return of capital gain and impermanent loss is halved by the price impact.

(a)



(b)



(c)

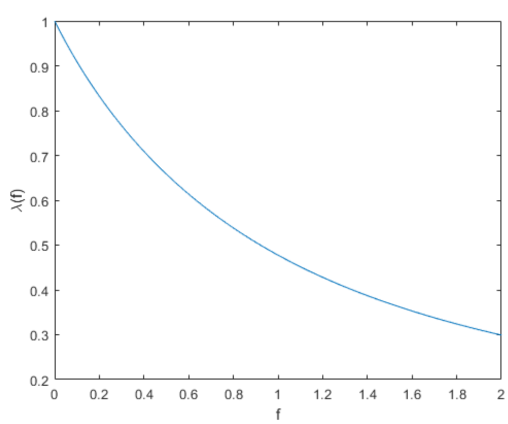


Figure 4: Liquidity and Offered Farm Yield

In this figure, we show the relation between a yield farm's offered yield and its aggregate liquidity. The x-axis corresponds to the logarithm of size of liquidity in the yield farm in units of \$1 million. The y-axis corresponds to the logarithm of one plus the annualized offered farm yield measured in decimal units. (For example, 50% of the annualized farm yield is 0.5 in decimal units.) The blue dots are observations measured at a daily frequency. The red dashed line plots the best linear fit obtained by regressing the logarithm of (1 + annualized offered farm yield) on the logarithm of the size of liquidity in the yield farm.

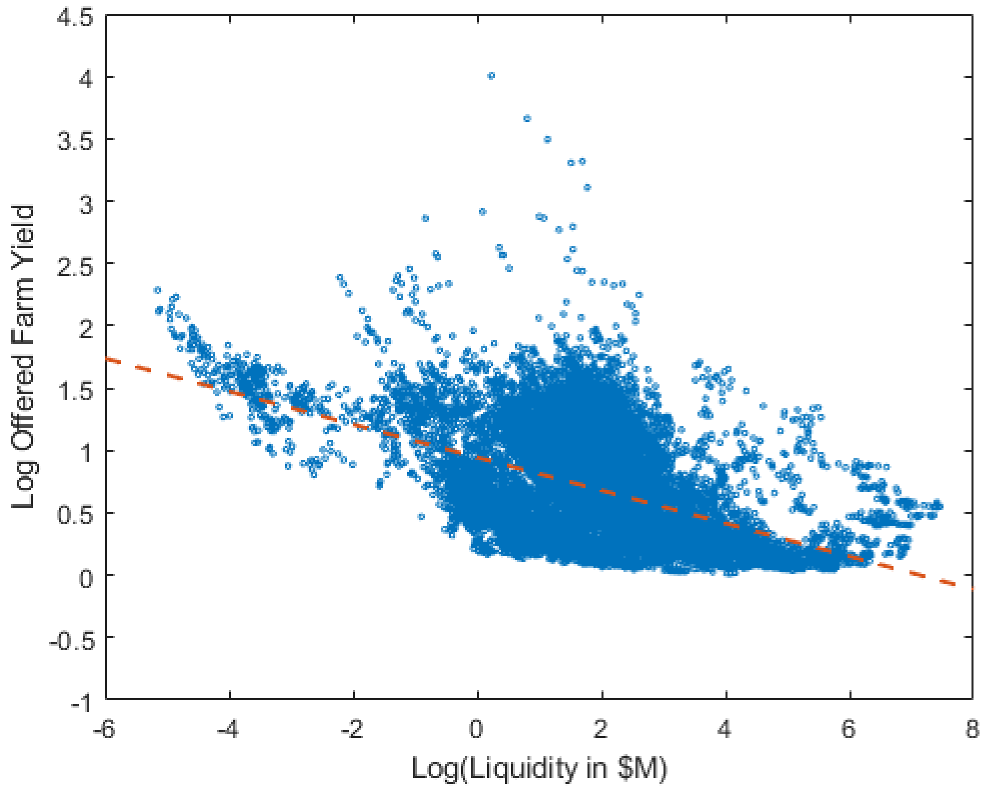
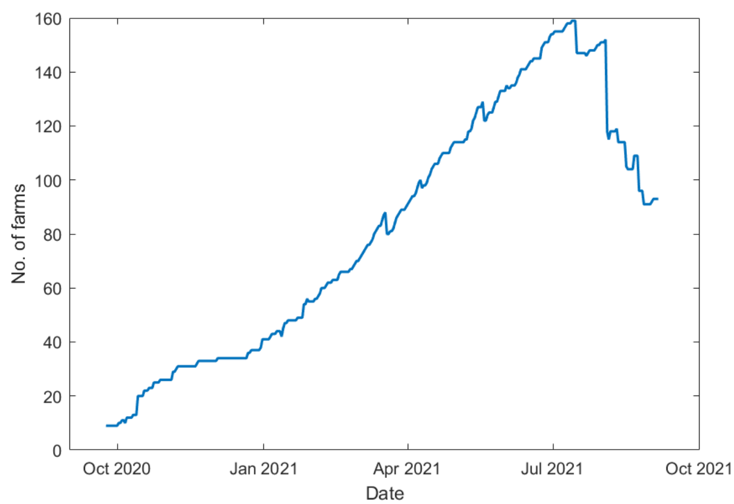


Figure 5: Number of Active Yield Farms and Total Value Locked in PancakeSwap

In this figure, we plot the number of active farms and Total Value Locked (TVL) in the sample period between September 23, 2020 (beginning of yield farming at PancakeSwap) and September 5, 2021, at a weekly frequency. In Panel (a), we provide the time series of active farms during our sample period. We define active farms as farms whose multipliers are larger than 0, implying that investors who stake LP tokens in these farms receive non-negative yields. In Panel (b), we plot TVL of active farms, or the amount of liquidity deposited for yield farming. The vertical axis is in millions of USD.

(a)



(b)

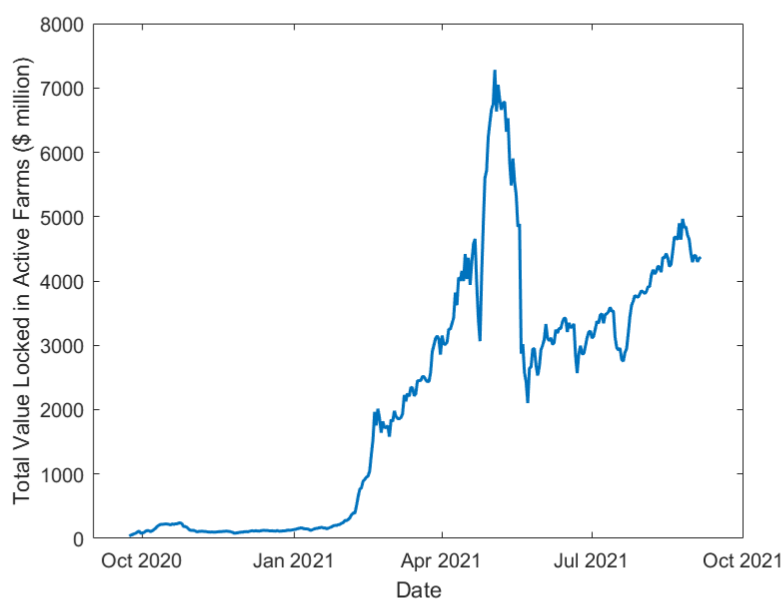


Figure 6: Duration of Yield Farms

In this figure, we plot a histogram of yield farm durations (in terms of days) during our sample period starting on October 20, 2021 and ending on September 5, 2021. The duration of a yield farm is defined as the number of days during which a farm's yield multiplier is positive. Thus, we consider a yield farm to be active/alive as long as staked LP tokens generate non-zero passive income for yield farmers.

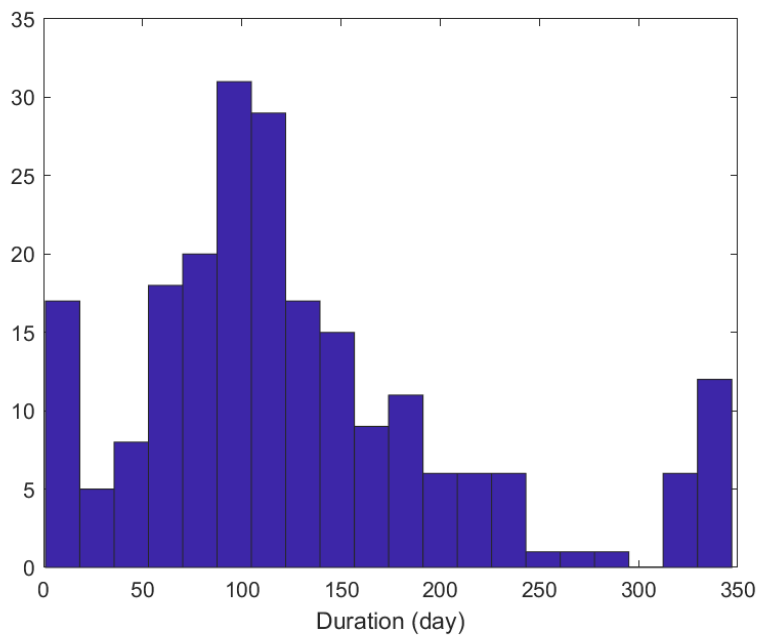
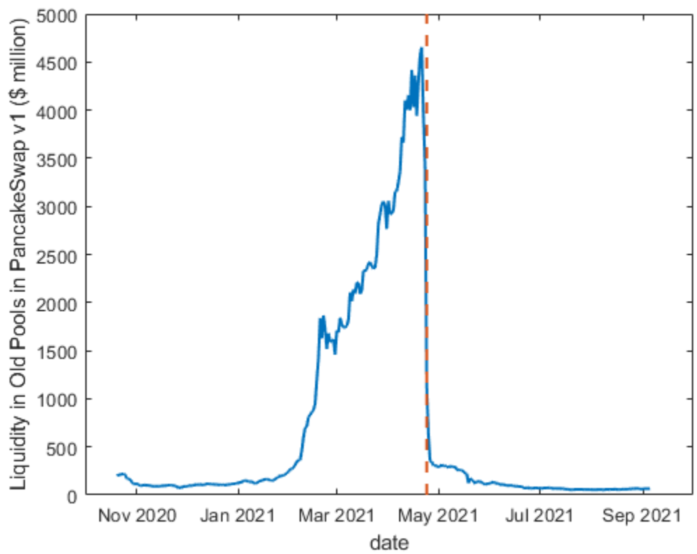


Figure 7: Remaining liquidity in PancakeSwap v1 after the launch of PancakeSwap v2

In this figure, we plot total value locked in liquidity pools of yield farms at PancakeSwap v1 whose new counterpart yield farms are available in PancakeSwap v2. On April 24, 2021, farms corresponding to liquidity pools in PancakeSwap v1 stopped providing farm yields. Instead, PancakeSwap encouraged farmers to move to corresponding counterpart farms available in PancakeSwap v2 so that the existing yield farmers could continue to earn farm yields. The blue lines in Panels (a) and (b) are total value locked in the liquidity pools whose new counterpart yield farms are available in PancakeSwap v2. The red dashed lines in Panels (a) and (b) indicate the date on which PancakeSwap v2 was launched. The sample period in Panel (a) is from October 20, 2020 to September 5, 2021, whereas Panel (b) focuses on the period since PancakeSwap v2 was launched.

(a)



(b)

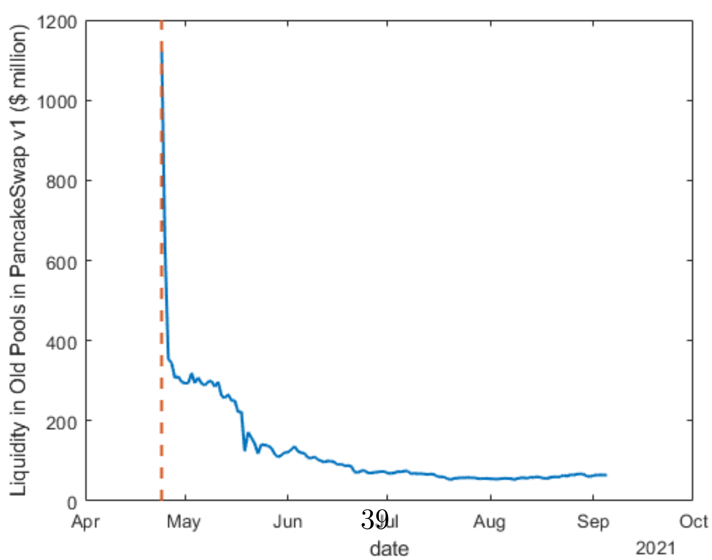


Figure 8: Staking Ratio of LP Tokens

In this figure, we plot the ratio of LP tokens staked in active yield farms listed in PancakeSwap, relative to the total number of LP tokens distributed as rewards for liquidity provision in the liquidity pools. Thus, the LP staking ratio is defined as the number of LP tokens of a liquidity pool staked in its corresponding farm, divided by the total number of outstanding LP tokens for the liquidity pool. The solid blue line indicates the median annualized offered farm yield. Dark and light shaded areas represent the interquartile range, as well as the 10th and 90th percentiles of the yield farm distribution, respectively.

Figure 9: Average Gas Fee to Enter and Exit a Yield Farming Position

In this figure, we compute the average gas fee paid by users on PancakeSwap (Panel (a)) and SushiSwap (Panel (b)) to enter (exit) a yield farming position on each day since the inception of the respective platform. For one round of yield farming, the total gas fee paid is the entry fee on the portfolio formation day, plus the exit fee on the last day of the holding period. For PancakeSwap, the average cost to enter (exit) over all days is \$1.39 (\$1.89). For SushiSwap, the average cost to enter (exit) over all days is \$109.12 (\$164.34).

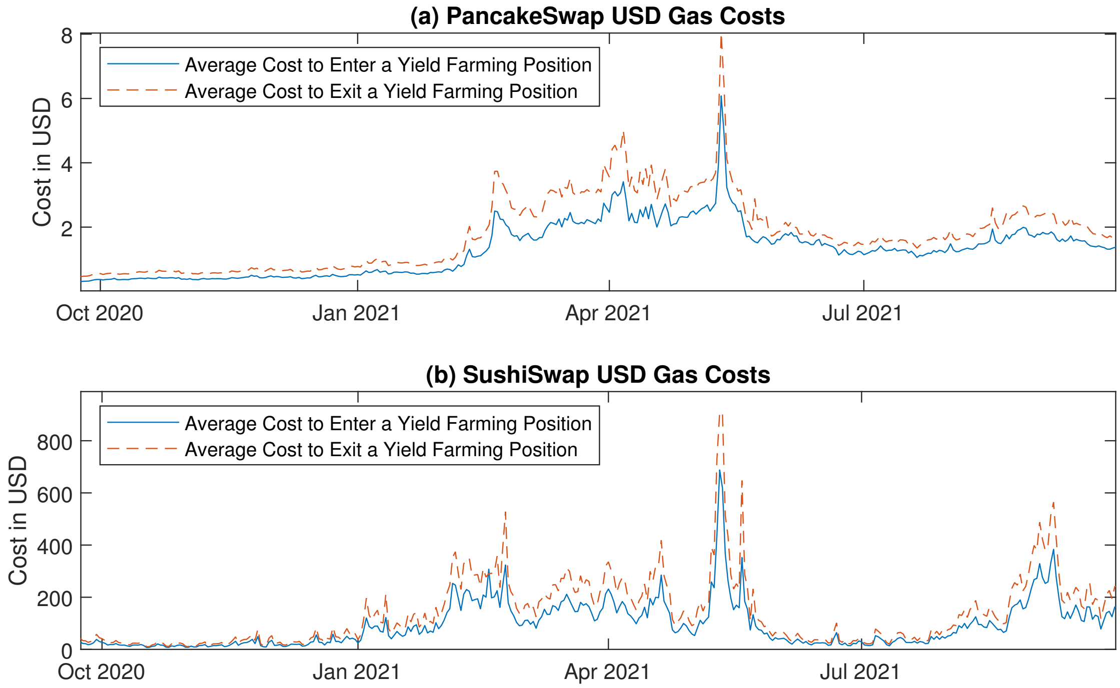
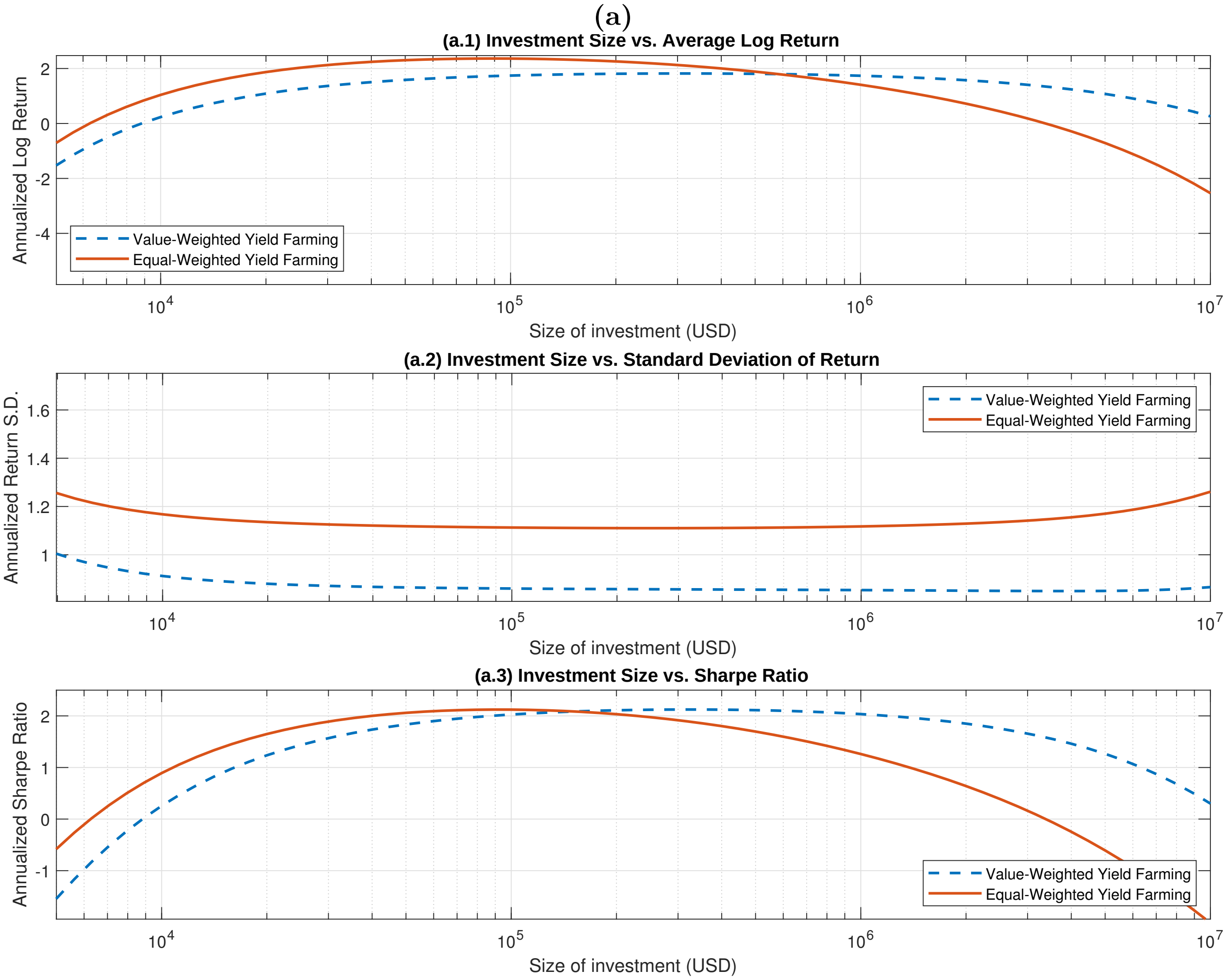


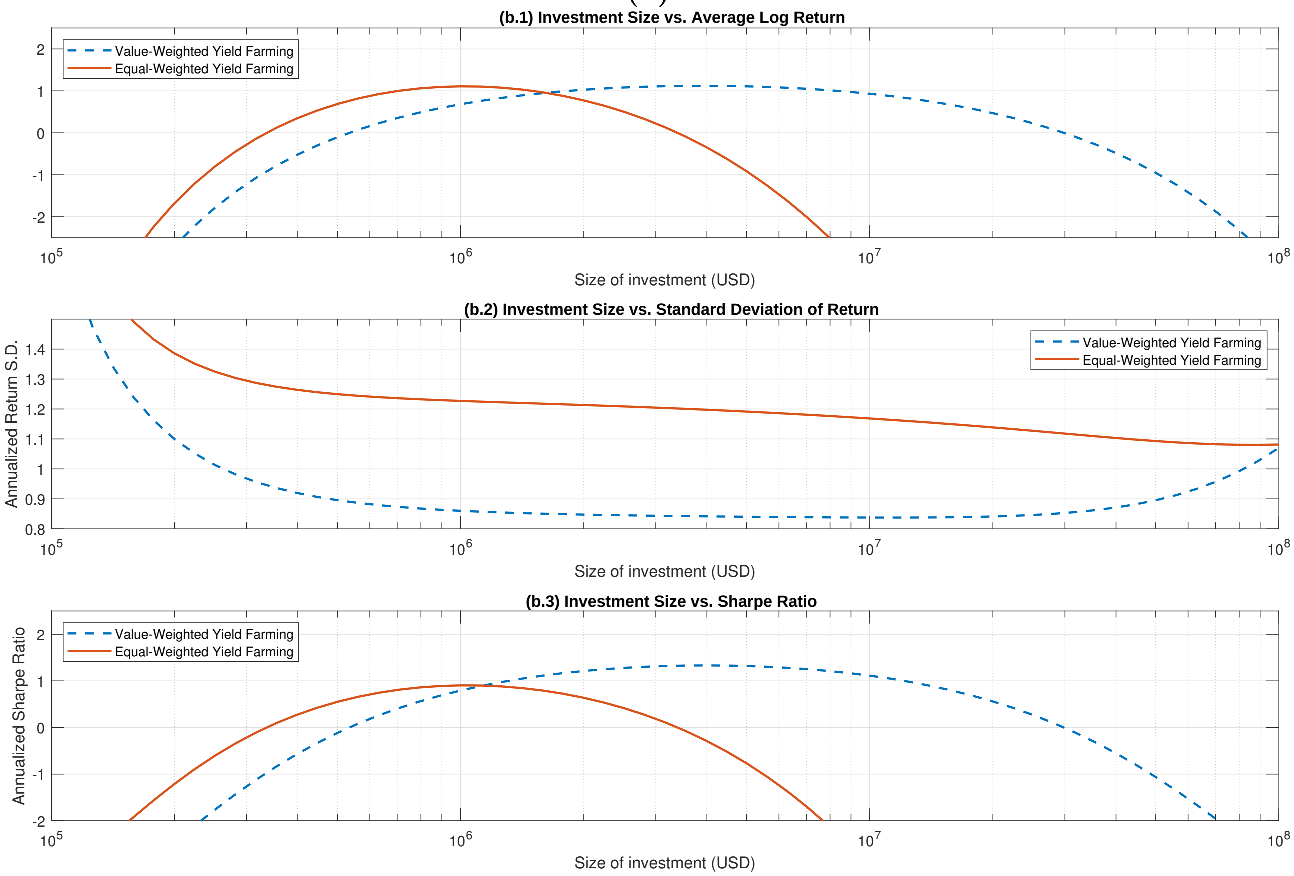
Figure 10: Investment Size and Investment Performance

In this figure, we plot the hypothetical performance of two diversified yield-farming investment strategies in PancakeSwap and SushiSwap as a function of investment size. In Panel (a) and Panel (b), we present the results regarding PancakeSwap and SushiSwap, respectively. The Value-Weighted (Equal-Weighted) investment strategy consists of a value-weighted (equal-weighted) portfolio of all yield farm returns, beginning on October 20, 2020 and re-balancing every seven days thereafter until September 5, 2021. In all figures, the x-axis is the investment size on a logarithmic scale. The y-axis in the first panel is the annualized average return. The y-axis in the second panel is the annualized standard deviation of returns. The y-axis in the third panel is the annualized Sharpe ratio of returns. More details for specific investment sizes can be found in Table 9.

(a)



(b)



(b.1) Investment Size vs. Average Log Return
Value-Weighted Yield Farming
Equal-Weighted Yield Farming
Size of investment (USD)
(b.2) Investment Size vs. Standard Deviation of Return
Annualized Return S.D.
Value-Weighted Yield Farming
Equal-Weighted Yield Farming
Size of investment (USD)
(b.3) Investment Size vs. Sharpe Ratio
Annualized Sharpe Ratio
Value-Weighted Yield Farming
Equal-Weighted Yield Farming
Size of investment (USD)

Figure 11: Relationship between Return on Yield Farming and Flow to Yield Farms

In this figure, we plot the correlation between log returns on yield farming and flows to a farm. Return on yield farming and flow to a farm are defined in Section 4. The x-axis is the week difference (w) between the timing of returns and flows. If w is positive, the timing of the flow is w weeks after that of the return. If w is negative, the timing of return is w weeks after that of the flow. The blue bar plots the correlation coefficients for the two variables and the red error bar plots the 95% confidence interval for the estimated correlation.

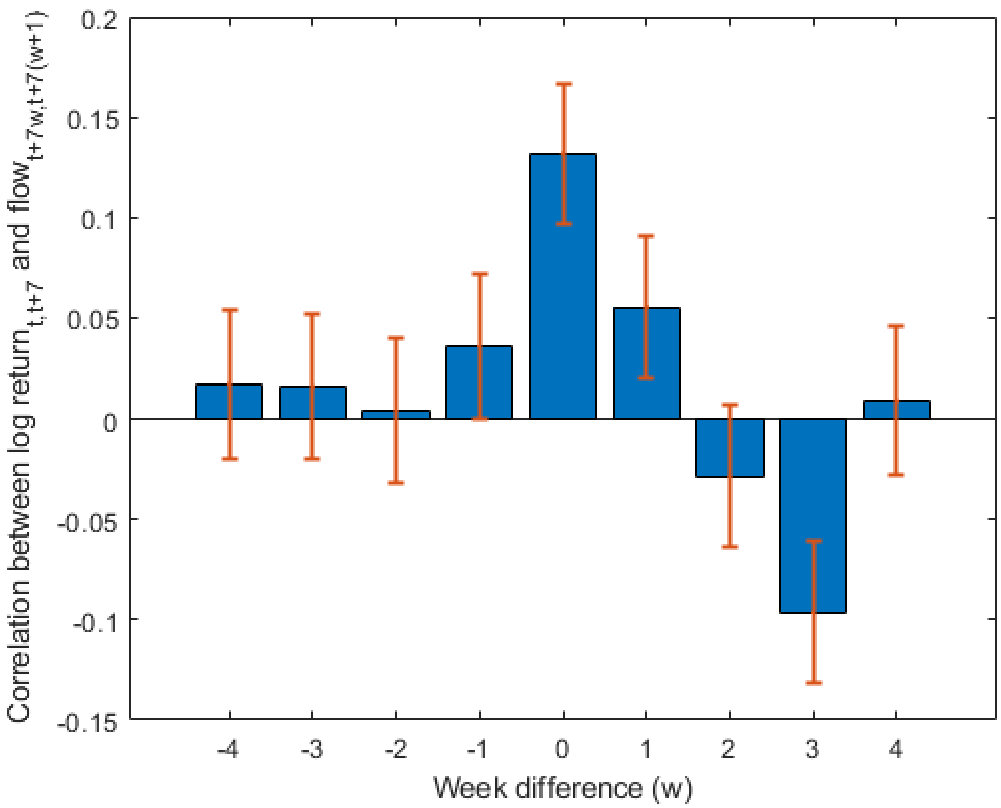


Table 1: Literature on Decentralized Finance and Decentralized Exchanges

This table summarizes a selection of key academic studies that focus on decentralized exchanges (DEXs) within the emerging ecosystem of decentralized finance. We indicate whether the study is primarily of empirical or theoretical nature, and list the decentralized platforms studied in each paper: Uniswap, SushiSwap, PancakeSwap. We also emphasize whether the study focuses on liquidity mining/provision and market making, strategic trading and hedging or yield farming.

	Theory vs. Empirical		**DEX**			**Activity**		
Study	**Theory**	**Empirical**	**Uniswap**	**SushiSwap**	**PancakeSwap**	**Liquidity Provision/ Market Making**	**Strategic Trading/ Hedging**	**Yield Farming**
Angeris, Kao, Chiang, Noyes, and Chitra (2019)	✓		✓			✓		
Aoyagi (2021)	✓		✓			✓		
Aoyagi and Ito (2021)	✓		✓			✓	✓	
Neuder, Rao, Moroz, and Parkes (2021)	✓		✓			✓	✓	
Park (2021)	✓		✓			✓	✓	
Lehar and Parlour (2021)	✓	✓	✓			✓	✓	
Han, Huang, and Zhong (2021)		✓	✓				✓	
Capponi and Jia (2021)	✓	✓	✓	✓			✓	
This study		✓			✓	✓		✓

Table 2: Top 10 Cryptocurrency Decentralized Exchanges

In this table, we report information regarding the 10 largest cryptocurrency decentralized exchanges in terms of daily trading volume as of October 9, 2021. For each exchange, we provide information on the daily trading volume (in $ million), the market share (in %), the number of markets at the exchange, the exchange type (swap, aggregator, order book, ...), whether spots or derivatives are traded on a DEX, and the month/year in which the exchange was launched. Source: https://coinmarketcap.com/rankings/exchanges/dex/.

Rank	**DEX**	**Daily Volume ($ million)**	**Mkt Share (%)**	**# Markets**	**Type**	**Spot /Derivatives**	**Launch Date**
1	dYdX	$1,756.41	25.05%	13	Orderbook	Derivatives	Apr 2019
2	PancakeSwap (V2)	$1,185.34	16.90%	1667	Swap	Spot	Apr 2021
3	Uniswap (V3)	$789.82	11.26%	627	Swap	Spot	May 2021
4	1inch Liquidity Protocol	$515.69	7.35%	26	Swap	Spot	Dec 2020
5	Uniswap (V2)	$287.57	4.10%	1556	Swap	Spot	Nov 2018
6	Sushiswap	$278.78	3.98%	387	Swap	Spot	Sep 2020
7	Honeyswap	$220.18	3.14%	66	Swap	Spot	Jul 2020
8	MDEX	$206.81	2.95%	140	Swap	Spot	Jan 2021
9	QuickSwap	$96.52	1.38%	330	Swap	Spot	Oct 2020
10	Raydium	$93.89	1.34%	112	Swap	Spot	Feb 2021

Table 3: Top 10 Yield Farms in PancakeSwap

In this table, we report information regarding top 10 farms in terms of total value locked (TVL) as of the end of our sample period, September 5, 2021. For each farm defined by a unique cryptocurrency pair, we provide information on the start date of a farm, annualized offered farm yield (in %), and total value locked (TVL, in $ million).

Farm Rank	**Cryptocurrency Pairs**	**Start Date**	**TVL ($ million)**	**Offered Farm Yield (%)**
1	CAKE-BNB	23-Sep-2020	931.72	34.90%
2	BUSD-BNB	23-Sep-2020	506.74	18.06%
3	USDT-BUSD	1-Oct-2020	285.28	7.15%
4	USDT-BNB	13-Oct-2020	278.04	17.51%
5	ETH-BNB	6-Oct-2020	250.73	8.14%
6	BTCB-BNB	6-Oct-2020	169.60	12.07%
7	USDC-BUSD	12-Jan-2021	140.96	5.82%
8	MBOX-BNB	8-Jun-2021	122.68	6.62%
9	BTCB-BUSD	29-Apr-2021	114.71	17.69%
10	USDC-USDT	28-Jun-2021	103.29	7.86%

Table 4: Summary Statistics

In this table, we report summary statistics on the return characteristics from yield farming and alternative investment strategies. The sample period is October 20, 2020 to September 5, 2021. The sample includes 219 unique liquidity pools associated with 219 unique yield farms. In Panel A, we report the cross-sectional average daily mean ($Mean$), median ($Median$), 25th ($p25$) and 75th ($p75$) percentiles of the log return distribution and the corresponding standard deviation (SD), skewness ($Skew$), kurtosis ($Kurt$), the first order autocorrelation coefficient ($AC1$), the number of time series ($\#TS$) and the average number of observations for each time series (OBS). In Panel B, we report the same information aggregated at a weekly frequency starting from October 20, 2020. All return-based statistics are annualized.

Panel A: Daily										
Variable	**Mean**	**SD**	**p25**	**Median**	**p75**	**Skew**	**Kurt**	**AC1**	**#TS**	**OBS**
Yield Farming Related Strategy										
Yield Farming	1.4770	1.2589	-10.7408	1.9316	13.8564	-0.3061	7.6484	-0.0999	219	129.4749
Liquidity Mining	0.1944	1.2869	-12.2038	0.5753	12.6476	-0.3148	7.7429	-0.0974	219	129.4749
Buy and Hold (Capital Gain)	0.4910	1.2770	-11.7385	0.5155	12.2615	-0.0900	7.8647	-0.0944	219	129.4749
Benchmark Strategy										
Bitcoin	1.6872	0.8184	-6.6577	0.9842	10.2659	-0.0602	4.5463	-0.0723	1	321
Ethereum	2.6641	1.1056	-9.3253	3.3788	15.5856	-0.5416	7.4258	-0.0893	1	321
MVIS 10 index	2.2795	1.0574	-7.8283	2.9056	15.0479	-0.7909	7.4093	-0.1282	1	321
MVIS 100 index	1.9048	0.9186	-6.9797	1.6245	12.8309	-0.6778	6.5268	-0.1336	1	321
Crypto Market Return	1.8627	0.9469	-5.7472	3.7891	11.9020	-1.4206	10.9296	-0.1654	1	321
S&P 500 index	0.3170	0.1337	-0.7547	0.3198	1.6078	-0.5792	4.9087	-0.0679	1	221

Panel B: Weekly										
Variable	**Mean**	**SD**	**p25**	**Median**	**p75**	**Skew**	**Kurt**	**AC1**	**#TS**	**OBS**
Yield Farming Related Strategy										
Yield Farming	1.6976	1.1884	-2.9113	1.9471	6.6150	-0.1660	3.7229	0.0407	219	20.3881
Liquidity Mining	0.0630	1.2472	-4.7842	0.5348	5.2841	-0.2664	3.7847	0.0317	219	20.3881
Buy and Hold (Capital Gain)	0.3759	1.2015	-4.1795	0.4920	5.0737	-0.1368	4.0227	0.0230	219	20.3881
Benchmark Strategy										
Bitcoin	1.6072	0.8039	-1.5417	1.9565	5.0711	-0.1920	3.6945	0.0506	1	45
Ethereum	2.4798	0.9592	-1.2047	1.8896	7.6812	-0.0942	3.4519	0.1703	1	45
MVIS 10 index	2.1920	0.8930	-1.0153	2.1312	5.5719	-0.5030	3.9603	0.1024	1	45
MVIS 100 index	1.8275	0.7417	-0.4450	2.0685	4.1985	-0.7626	4.4258	0.1546	1	45
Crypto Market Return	1.7272	0.7913	-1.5302	2.5532	5.1645	-1.0549	5.0982	0.2006	1	45
S&P 500 index	0.3230	0.1351	-0.3794	0.3520	0.6915	0.7631	4.6875	-0.4165	1	45

Table 5: Return Decomposition

In this table, we decompose each return series into the contributions arising from (a) capital gains, (b) impermanent losses, (c) trading fees, and (d) farm yields. The sample period is October 20th, 2020 to September 5th, 2021. In Panel A, we report summary statistics on the return characteristics for each component. We report the cross-sectional average daily mean log return (Ret) median ($Median$), 25th ($p25$) and 75th ($p75$) percentiles of the log return distribution and the corresponding standard deviation (SD), skewness ($Skew$), kurtosis ($Kurt$), the first order autocorrelation coefficient ($AC1$), and the average number of observations for each time series (OBS). We also report the same information sorted by terciles in terms of average in-sample offered yield. In Panel B, we report the same information aggregated at a weekly frequency starting from October 20, 2020. All return-based statistics are annualized.

Panel A: Daily **Component**	**Mean**	**SD**	**p25**	**Median**	**p75**	**Skew**	**Kurt**	**AC1**	**OBS**
Full Sample									
Capital Gains	0.3829	1.2473	-11.8466	0.5121	12.3606	-0.0750	7.7172	-0.1022	129.4749
Impermanent Loss	-0.4108	0.0767	-11.8466	-0.0884	-0.0220	-5.5866	49.4831	0.0951	129.4749
Trading Fees	0.1136	0.0072	0.0475	0.0768	0.1348	2.9542	18.4164	0.4701	129.4749
Farm Yields	1.5392	0.0308	1.1351	1.4704	1.8573	0.9113	5.4694	0.7625	129.4749
Tercile 1									
Capital Gains	1.5524	1.0403	-8.2413	1.3081	11.1555	0.0160	8.4568	-0.1375	176.9863
Impermanent Loss	-0.2530	0.0473	-8.2413	-0.0493	-0.0107	-6.6853	71.5797	0.1203	176.9863
Trading Fees	0.1137	0.0093	0.0448	0.0722	0.1318	3.0882	21.4939	0.5547	176.9863
Farm Yields	0.5298	0.0162	0.3268	0.4561	0.6560	1.5669	8.4307	0.8110	176.9863
Tercile 2									
Capital Gains	0.4381	1.2154	-11.0216	0.3898	11.6992	-0.0009	8.0638	-0.0888	113.2055
Impermanent Loss	-0.4002	0.0654	-11.0216	-0.0807	-0.0187	-5.3846	42.6768	0.1089	113.2055
Trading Fees	0.1144	0.0066	0.0457	0.0753	0.1347	3.1567	19.7484	0.4815	113.2055
Farm Yields	1.3681	0.0354	0.9040	1.2894	1.7581	0.6009	3.8779	0.8166	113.2055
Tercile 3									
Capital Gains	-0.8418	1.4861	-16.2769	-0.1616	14.2270	-0.2401	6.6310	-0.0805	89.2329
Impermanent Loss	-0.5792	0.1174	-16.2769	-0.1353	-0.0365	-4.6900	34.1930	0.0563	89.2329
Trading Fees	0.1127	0.0056	0.0519	0.0828	0.1378	2.6308	14.1237	0.3759	89.2329
Farm Yields	2.7196	0.0409	2.1746	2.6657	3.1579	0.5660	4.0995	0.6598	89.2329

Panel B: Weekly **Component**	**Mean**	**SD**	**p25**	**Median**	**p75**	**Skew**	**Kurt**	**AC1**	**OBS**
Full Sample									
Capital Gains	0.3461	1.2002	-4.1780	0.4610	5.0141	-0.1377	4.0485	0.0272	20.388
Impermanent Loss	-0.5088	0.1371	-4.1780	-0.1405	-0.0348	-2.2584	8.7455	0.0195	20.388
Trading Fees	0.1096	0.0147	0.0478	0.0743	0.1335	1.7011	6.0237	0.3370	20.388
Farm Yields	1.3756	0.0821	0.9822	1.2894	1.6796	0.6626	3.9011	0.5046	20.388
Tercile 1									
Capital Gains	1.5212	1.0023	-2.5983	1.4304	5.6450	0.0347	4.2930	0.0100	27.299
Impermanent Loss	-0.3026	0.0900	-2.5983	-0.0843	-0.0199	-2.7355	12.4430	0.0083	27.299
Trading Fees	0.1143	0.0142	0.0497	0.0804	0.1403	1.5098	5.4746	0.4405	27.299
Farm Yields	0.5120	0.0414	0.3160	0.4306	0.6290	1.1461	4.7528	0.5899	27.299
Tercile 2									
Capital Gains	0.1257	1.1676	-4.0669	0.4520	4.3535	-0.1662	4.2069	-0.0238	18.045
Impermanent Loss	-0.5512	0.1567	-4.0669	-0.1402	-0.0351	-2.2074	8.0561	-0.0019	18.045
Trading Fees	0.1279	0.0165	0.0592	0.0857	0.1609	1.8292	6.4456	0.2905	18.045
Farm Yields	1.3377	0.0938	0.9119	1.2394	1.6932	0.4193	3.6178	0.5361	18.045
Tercile 3									
Capital Gains	-0.6087	1.4307	-5.8688	-0.4995	5.0439	-0.2816	3.6458	0.0953	15.821
Impermanent Loss	-0.6726	0.1645	-5.8688	-0.1969	-0.0494	-1.8322	5.7375	0.0520	15.821
Trading Fees	0.0865	0.0132	0.0345	0.0569	0.0993	1.7691	6.1683	0.2789	15.821
Farm Yields	2.2773	0.111	1.7187	2.1982	2.7166	0.4224	3.3329	0.3878	15.821

Table 6: Variance Decomposition

In this table, we report the proportion of variation in the aggregate return series explained by each of the four return components arising from (a) capital gains, (b) impermanent losses, (c) trading fees, and (d) farm yields. For the variance decomposition, we split the covariance terms equally for each component. We report the variance decomposition for the overall population, and for each tercile of the population sorted by their average offered yield throughout the sample period. The sample period is October 20, 2020 to September 5, 2021.

Panel A: Daily				
Component	**Full Sample**	**Tercile 1**	**Tercile 2**	**Tercile 3**
Capital Gains	0.9995	1.0077	1.0073	0.9836
Impermanent Loss	-0.0059	-0.0132	-0.0137	0.0091
Trading Fees	0.0011	0.0019	0.0011	0.0003
Farm Yields	0.0053	0.0036	0.0053	0.0071
Panel B: Weekly				
Component	**Full Sample**	**Tercile 1**	**Tercile 2**	**Tercile 3**
Capital Gains	1.0055	1.0073	1.0142	0.9949
Impermanent Loss	-0.0141	-0.0304	-0.0159	0.0039
Trading Fees	0.0072	0.0191	0.0016	0.0010
Farm Yields	0.0014	0.0041	0.0001	0.0002

Table 7: Lack of Sophistication of Yield Farmers

In this table, we report the statistics that describe behaviors of yield farmers. The presented statistics are all farmer-level variables. In Panel A, we present summary statistics of yield farmers. *No. Farms* is the number of farms in which an yield farmer invests. *Staked Balance* is the dollar value of LP tokens staked in farms. *LP Balance* is the dollar value of LP tokens. *Efficiency* is average of efficiency of each farm. Efficiency is defined as the length of time for which the user has staked his/her LP tokens in a farm divided by the length of the time for which the user has kept the liquidity in a corresponding liquidity pool. *Offered Farm Yield* is time-weighted average of the offered yield at the beginning of the holding period. In Panel B, we decompose the yield farmers into two groups depending on the *Efficiency* using 0.98 as a threshold. In Panel C, we decompose the yield farmers into quintiles by *LP Balance*.

Panel A: Yield Farmers

Variables	**Mean**	**SD**	**p25**	**Median**	**p75**	**OBS**
No. Farms	1.81	1.52	1.00	1.00	2.00	207,699
Staked Balance ($)	6,539.36	201,986.99	33.95	173.26	823.89	207,699
LP Balance ($)	6,959.62	203,816.10	61.88	222.05	954.66	207,699
Efficiency	0.82	0.35	0.95	1.00	1.00	207,699
Offered Farm Yield	1.53	1.35	0.27	1.24	2.59	207,699

Panel B: Yield Farmers by Efficiency

	No. Farms	**Staked Balance ($)**	**LP Balance ($)**	**Efficiency**	**Offered Farm Yield**	**OBS**
Efficiency≤0.98						
Mean	1.59	1,695.34	3,099.74	0.40	1.40	60,514
(SD)	(1.36)	(37,885.67)	(62,729.03)	(0.41)	(1.48)	
Efficiency>0.98						
Mean	1.90	8,530.95	8,546.58	0.998	1.58	147.185
(SD)	(1.57)	(238,682.27)	(238,734.04)	(0.004)	(1.29)	

Panel C: Yield Farmers by LP Balance.

	No. Farms	**Staked Balance ($)**	**LP Balance ($)**	**Efficiency**	**Offered Farm Yield**	**OBS**
Quintile 1						
Mean	1.446	17.938	25.553	0.672	1.530	41540
S.D.	*(0.990)*	*(14.335)*	*(10.581)*	*(0.437)*	*(1.500)*	
Quintile 2						
Mean	1.592	68.277	84.421	0.797	1.610	41540
S.D.	*(1.181)*	*(38.069)*	*(24.726)*	*(0.368)*	*(1.374)*	
Quintile 3						
Mean	1.829	204.396	236.434	0.855	1.614	41539
S.D.	*(1.493)*	*(99.663)*	*(71.232)*	*(0.318)*	*(1.313)*	
Quintile 4						
Mean	1.999	660.956	742.677	0.883	1.581	41540
S.D.	*(1.676)*	*(324.585)*	*(260.844)*	*(0.290)*	*(1.298)*	
Quintile 5						
Mean	2.198	31745.103	33708.856	0.917	1.310	41540
S.D.	*(1.941)*	*(450779.453)*	*(454767.188)*	*(0.248)*	*(1.235)*	

Table 8: Returns from Yield Farming Portfolios

This table reports the summary statistics for percentage excess returns from yield farming investment strategies. We take the perspective of a U.S. investor and report all information from the perspective of an initial USD investment. Excess returns are computed relative to the three-month U.S. Treasury bill secondary market rate source from the Federal Reserve Bank of St.Louis. All returns are value-weighted using the pools' aggregate liquidity as weighting factors. The column (OBS) reports the number of observations. We report the mean return ($Mean$), the standard deviation, 25th percentile, median, 75th percentile, skewness, and kurtosis of the yield farming strategies, as well as the serial correlation, the Sharpe ratio, the alpha from a three factor model based on the work of Liu, Tsyvinski, and Wu (2019), and the t-statistic for alpha from the three-factor regressions. The sample period is October 20, 2020 to September 5, 2021. All return-based statistics are annualized. Because we report excess returns and alphas as annualized log returns, the mean return and alpha can be lower than -1, unlike arithmetic returns.

Panel A: Daily												
Strategy	**Mean**	**SD**	**p25**	**Median**	**p75**	**Skew**	**Kurt**	**AC1**	**SR**	**α**	**t-stat of α**	**OBS**
Yield Farming Related Strategy												
Yield Farming	2.3712	0.8194	-4.7524	2.5652	10.8074	-0.2228	9.8805	-0.0969	2.8938	1.3778	1.8653	321
Liquidity Mining	1.8598	0.8164	-5.2658	1.8591	10.3550	-0.2635	9.8465	-0.1028	2.2782	0.4454	0.7994	321
Buy and Hold (Capital Gain)	1.9628	0.8235	-5.3285	1.9200	10.4074	-0.0110	9.9887	-0.0979	2.3835	0.6137	1.0024	321
Benchmark Strategy												
Bitcoin	1.6872	0.8184	-6.6577	0.9842	10.2659	-0.0602	4.5463	-0.0723	2.0617	-0.2851	-0.6508	321
Ethereum	2.6641	1.1056	-9.3253	3.3788	15.5856	-0.5416	7.4258	-0.0893	2.4096	0.1948	0.2920	321
MVIS 10 index	2.2795	1.0574	-7.8283	2.9056	15.0479	-0.7909	7.4093	-0.1282	2.1557			321
MVIS 100 index	1.9048	0.9186	-6.9797	1.6245	12.8309	-0.6778	6.5268	-0.1336	2.0736			321
Crypto Market Return	1.8627	0.9469	-5.7472	3.7891	11.9020	-1.4206	10.9296	-0.1654	1.9672			321
S&P 500 index	0.3170	0.1337	-0.7547	0.3198	1.6078	-0.5792	4.9087	-0.0679	2.3707			221
Panel B: Weekly												
Strategy	**Mean**	**SD**	**p25**	**Median**	**p75**	**Skew**	**Kurt**	**AC1**	**SR**	**α**	**t-stat of α**	**OBS**
Yield Farming Related Strategy												
Yield Farming	2.4294	0.8182	-0.7666	1.8066	4.4608	0.2980	3.6981	0.2875	2.9692	1.3125	1.6202	45
Liquidity Mining	1.9008	0.8048	-1.4499	1.4603	4.1851	0.2101	3.6353	0.2661	2.3620	0.3867	0.6565	45
Buy and Hold (Capital Gain)	2.0114	0.8322	-1.2989	1.5381	4.2499	0.3494	3.8112	0.2664	2.4170	0.4965	0.7489	45
Benchmark Strategy												
Bitcoin	1.6072	0.8039	-1.5417	1.9565	5.0711	-0.1920	3.6945	0.0506	1.9993	0.0256	0.0405	45
Ethereum	2.4798	0.9592	-1.2047	1.8896	7.6812	-0.0942	3.4519	0.1703	2.5853	0.9998	0.9839	45
MVIS 10 index	2.1920	0.8930	-1.0153	2.1312	5.5719	-0.5030	3.9603	0.1024	2.4547			45
MVIS 100 index	1.8275	0.7417	-0.4450	2.0685	4.1985	-0.7626	4.4258	0.1546	2.4639			45
Crypto Market Return	1.7272	0.7913	-1.5302	2.5532	5.1645	-1.0549	5.0982	0.2006	2.1829			45
S&P 500 index	0.3230	0.1351	-0.3794	0.3520	0.6915	0.7631	4.6875	-0.4165	2.3919			45

Table 9: Impact of Trading Frictions on Returns from Yield Farming Portfolios

This table reports the summary statistics for percentage excess returns from yield farming investment strategies, accounting for trading costs and price impact. We take the perspective of a U.S. investor and report all information from the perspective of an initial USD investment. Excess returns are computed relative to the three-month U.S. Treasury bill secondary market rate sourced from the Federal Reserve Bank of St.Louis. All returns are value-weighted using the pools' aggregate liquidity as weighting factors. The column (OBS) reports the number of observations. We report the mean return ($Mean$), the standard deviation, skewness, and kurtosis of the yield farming strategies. We also report the Sharpe ratio (SR), information ratio (IR), the alpha from a three factor model based on the work of Liu, Tsyvinski, and Wu (2019), and the t-statistic for alpha from the three-factor regressions. The sample period is October 20, 2020 to September 5, 2021. All return-based statistics are annualized. Because we report excess returns and alphas as annualized log returns, the mean return and alpha can be lower than -1, unlike arithmetic returns.

Panel A: Daily							
Strategy	**Mean**	**SD**	**SR**	**IR**	α	**t-stat of α**	**OBS**
Frictionless benchmark							
Yield Farming (Full Sample)	2.3712	0.8194	2.8938	2.0745	0.8624	1.8653	321
Yield Farming (Tercile 1)	2.0915	0.7198	2.9055	2.0981	0.7668	1.8865	321
Yield Farming (Tercile 2)	3.2008	1.1479	2.7885	1.8959	0.1006	1.7047	321
Yield Farming (Tercile 3)	3.0816	1.1921	2.5851	1.1550	0.7860	1.0385	321
Gas fee, Trading fee & Price impact ($5,000)							
Yield Farming (Full Sample)	-26.4169	1.3840	-19.0869	-25.0460	-27.9663	-22.5199	321
Yield Farming (Tercile 1)	-7.7689	0.8175	-9.5034	-18.5515	-9.1026	-16.6804	321
Yield Farming (Tercile 2)	-6.7097	1.2287	-5.4610	-12.8857	-8.8066	-11.5860	321
Yield Farming (Tercile 3)	-6.8976	1.2708	-5.4276	-11.7501	-9.1889	-10.5650	321
Gas fee, Trading fee & Price impact ($10,000)							
Yield Farming (Full Sample)	-13.5242	1.0002	-13.5209	-23.0259	-15.0443	-20.7035	321
Yield Farming (Tercile 1)	-3.7343	0.7526	-4.9621	-12.5993	-5.0622	-11.3286	321
Yield Farming (Tercile 2)	-2.6828	1.1772	-2.2790	-7.7721	-4.7775	-6.9883	321
Yield Farming (Tercile 3)	-2.9473	1.2199	-2.4161	-7.2670	-5.2353	-6.5341	321
Gas fee, Trading fee & Price impact ($100,000)							
Yield Farming (Full Sample)	-2.6171	0.8253	-3.1710	-9.7941	-4.1261	-8.8063	321
Yield Farming (Tercile 1)	-0.3128	0.7212	-0.4337	-4.4779	-1.6416	-4.0262	321
Yield Farming (Tercile 2)	0.4327	1.1480	0.3769	-2.8528	-1.6692	-2.5651	321
Yield Farming (Tercile 3)	-1.0320	1.1933	-0.8648	-4.8001	-3.3647	-4.3160	321
Gas fee, Trading fee & Price impact ($1,000,000)							
Yield Farming (Full Sample)	-2.6072	0.8204	-3.1778	-9.8129	-4.1454	-8.8232	321
Yield Farming (Tercile 1)	-1.5290	0.7270	-2.1032	-7.6089	-2.9082	-6.8415	321
Yield Farming (Tercile 2)	-3.8629	1.1715	-3.2975	-9.2186	-6.0497	-8.2888	321
Yield Farming (Tercile 3)	-16.3170	1.6737	-9.7489	-13.7716	-19.1047	-12.3826	321

Panel B: Weekly							
Strategy	**Mean**	**SD**	**SR**	**IR**	α	t**-stat of** α	**OBS**
Frictionless benchmark							
Yield Farming (Full Sample)	2.4294	0.8182	2.9692	1.9616	0.8345	1.6202	45
Yield Farming (Tercile 1)	2.2181	0.7227	3.0691	2.0365	0.7761	1.6821	45
Yield Farming (Tercile 2)	2.9351	1.1181	2.6250	1.5503	0.8426	1.2805	45
Yield Farming (Tercile 3)	3.4141	1.2372	2.7594	1.5633	1.1650	1.2912	45
Gas fee, Trading fee & Price impact ($5,000)							
Yield Farming (Full Sample)	-1.5554	0.9705	-1.6027	-6.2233	-3.4437	-5.1404	45
Yield Farming (Tercile 1)	0.8462	0.7563	1.1188	-1.7431	-0.6884	-1.4398	45
Yield Farming (Tercile 2)	1.5449	1.1528	1.3401	-1.1207	-0.6338	-0.9257	45
Yield Farming (Tercile 3)	1.9968	1.2713	1.5707	-0.4265	-0.3298	-0.3523	45
Gas fee, Trading fee & Price impact ($10,000)							
Yield Farming (Full Sample)	0.2140	0.8754	0.2444	-3.3110	-1.5210	-2.7348	45
Yield Farming (Tercile 1)	1.4016	0.7371	1.9016	-0.2204	-0.0847	-0.1821	45
Yield Farming (Tercile 2)	2.0990	1.1339	1.8512	-0.0569	-0.0315	-0.0470	45
Yield Farming (Tercile 3)	2.5397	1.2526	2.0275	0.3447	0.2629	0.2848	45
Gas fee, Trading fee & Price impact ($100,000)							
Yield Farming (Full Sample)	1.7146	0.8216	2.0867	0.2628	0.1123	0.2171	45
Yield Farming (Tercile 1)	1.8715	0.7231	2.5883	1.1305	0.4315	0.9338	45
Yield Farming (Tercile 2)	2.5264	1.1179	2.2599	0.8114	0.4457	0.6702	45
Yield Farming (Tercile 3)	2.7842	1.2345	2.2553	0.7709	0.5837	0.6368	45
Gas fee, Trading fee & Price impact ($1,000,000)							
Yield Farming (Full Sample)	1.7076	0.8148	2.0956	0.3298	0.1414	0.2724	45
Yield Farming (Tercile 1)	1.6874	0.7169	2.3536	0.7577	0.2913	0.6258	45
Yield Farming (Tercile 2)	1.9163	1.1038	1.7360	-0.1502	-0.0838	-0.1241	45
Yield Farming (Tercile 3)	0.4974	1.2779	0.3892	-1.5780	-1.3805	-1.3034	45

Table 10: Flow Regressions

In this table, we report the results for the liquidity flow regressions. We regress the farm *Flow* over the next 7 days (a week) on *Offered Farm Yield*, past *Flow*, past *Log return* on yield farming, past *Log crypto MKT return*, and *Log (Liquidity)*. *Flow*, *Offered Farm Yield*, past *Log return*, and *Log (Liquidity)* are defined in Section 4. *Log crypto MKT return* is logarithm of the cryptocurrency market return, the construction process for which is explained in Section 5.4 in detail. The sample period is October 20, 2020 to September 5, 2021. Standard errors are clustered at the farm level.*,**, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
				$Flow_{t,t+7}$		
$Offered\ Farm\ Yield_t$	0.0149**		0.0149**	0.0973***		0.114***
	(0.0063)		(0.0072)	(0.0138)		(0.0156)
$Flow_{t-7,t}$		0.0528*	0.0449		0.0391	-0.0125
		(0.0284)	(0.0299)		(0.0288)	(0.0305)
$Flow_{t-14,t-7}$		0.0118	0.00543		0.00808	-0.0227
		(0.0248)	(0.0257)		(0.0259)	(0.0269)
$Flow_{t-21,t-14}$		-0.0136	-0.0182		-0.00856	-0.0222
		(0.0220)	(0.0228)		(0.0223)	(0.0226)
$Flow_{t-28,t-21}$		-0.0259	-0.0295*		-0.0297*	-0.0440**
		(0.0167)	(0.0173)		(0.0176)	(0.0179)
$Log\ return_{t-7,t}$		0.122***	0.110***		0.164***	0.0969***
		(0.0386)	(0.0370)		(0.0373)	(0.0359)
$Log\ return_{t-14,t-7}$		0.00505	-0.00241		0.0581	0.0337
		(0.0396)	(0.0396)		(0.0378)	(0.0368)
$Log\ return_{t-21,t-14}$		0.0109	0.000981		0.0672*	0.0304
		(0.0376)	(0.0381)		(0.0369)	(0.0360)
$Log\ return_{t-28,t-21}$		0.0224	0.0154		0.0721**	0.0443
		(0.0368)	(0.0373)		(0.0359)	(0.0372)
$Log\ crypto\ MKT\ return_{t-7,t}$		-0.120***	-0.106***		-0.160***	-0.0965***
		(0.0414)	(0.0390)		(0.0399)	(0.0368)
$Log\ crypto\ MKT\ return_{t-14,t-7}$		-0.00342	0.00564		-0.0545	-0.0416
		(0.0439)	(0.0441)		(0.0429)	(0.0415)
$Log\ crypto\ MKT\ return_{t-21,t-14}$		-0.00578	0.00317		-0.0673*	-0.0603
		(0.0386)	(0.0393)		(0.0389)	(0.0374)
$Log\ crypto\ MKT\ return_{t-28,t-21}$		-0.0212	-0.0156		-0.0782**	-0.0859**
		(0.0388)	(0.0390)		(0.0385)	(0.0371)
$Log(Liquidity_t)$		0.0005	0.0044		-0.0514***	-0.0527***
		(0.0028)	(0.0031)		(0.0071)	(0.0100)
Farm FE				✓	✓	✓
N	3016	3012	3012	3016	3012	3012
R^2	0.003	0.009	0.011	0.086	0.084	0.129
adj. R^2	0.002	0.005	0.006	0.026	0.019	0.067

Appendix

Figure A.1: UI of Yield Farms in PancakeSwap

In this figure, we provide a snapshot of user-interface environment of yield farms in PancakeSwap.

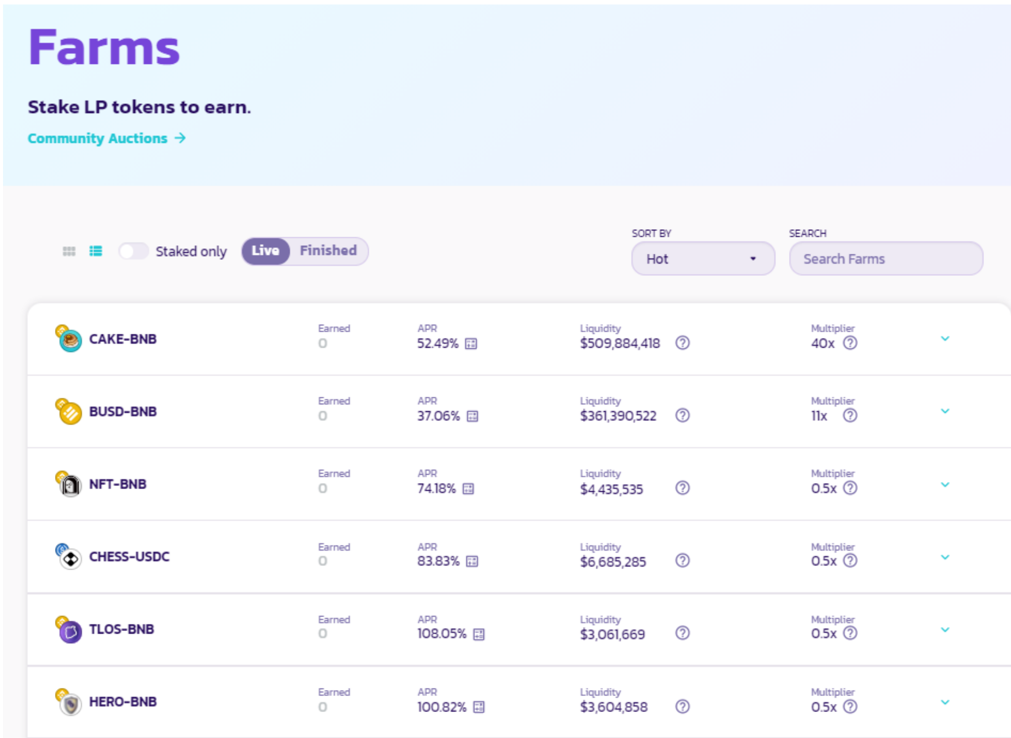


Figure A.2: Investment outcome from liquidity provision vs. Investment outcome from a buy-and-hold strategy

In this figure, we compare investment outcome from liquidity provision with investment outcome from a simple buy-and-hold strategy. A liquidity provider buys equal U.S. dollar amount of token A and token B of a liquidity pool at time t. y-axis is the ratio of investment outcome from a liquidity provision and investment outcome from a simple buy-and-hold strategy at time $t+h$ minus 1. x-axis is the growth of the ratio of prices of token A and token B between time t and $t+h$, i.e., $\frac{\rho_{t+h}}{\rho_t}$ where $\rho_t = \frac{P_t^A}{P_t^B}$

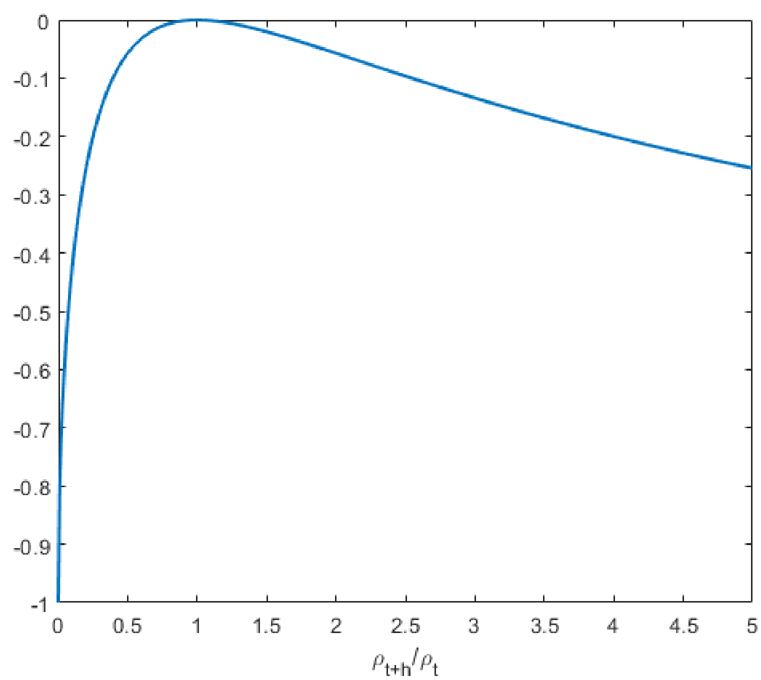


Figure A.3: Relation between Model-implied and Listed Offered Farm Yields

In this figure, we compare the offered farm yields calculated using Equation (8) on the y-axis to those listed on the PancakeSwap's homepage on the x-axis (https://pancakeswap.finance/farms). The listed farm yields are manually collected from Pancakeswap's web page at midnight Greenwich Meridian Time (GMT) on October 11, 2021. All values are reported in percentage points. The blue circles represent all observations and the red dashed line connects (0%,0%) and (300%,300%), i.e., a 45-degree line. A linear regression where we regress the calculated on the listed farm yields obtains an R^2 of 1.00 and an estimated regression line given by $\widehat{y}_t = 1.002 \times y_t - 0.001$.

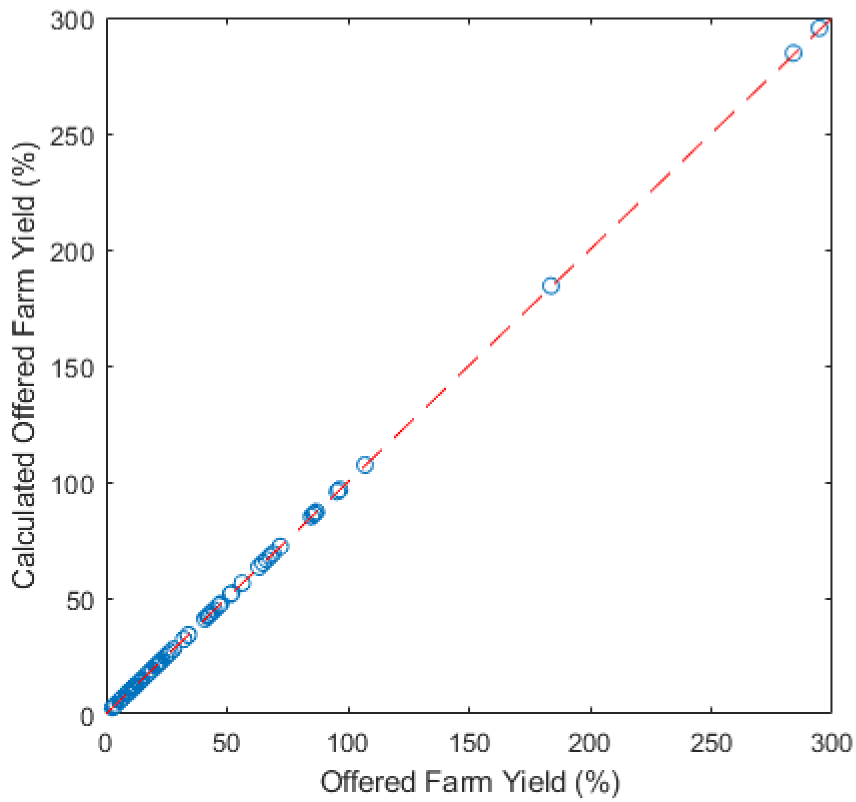


Table A.1: Chain of Transactions for Yield Farming Strategies

In this table, we itemize the individual transactions in a yield farming strategy. We explain how each step of the yield farming strategy can change the number of tokens in a liquidity pool and we describe three different types of transaction costs: gas fees, trading fees, and price impact. We refer to a hypothetical pair of cryptocurrency tokens A and B in a liquidity pool (LP) A/B.

Step	**Timing**	**Event**	**# Tokens A in LP for A/B**	**# Tokens B in LP for A/B**	**Trading Frictions**		
					Gas Fee	**Trading Fee**	**Price Impact**
1	t	Yield farming starts.	α_t^A	α_t^B			
2	t	The yield farmer buys Δ_t^A units of token A using a part of his/her fund, $I_t = fL_t$, using Δ_t^B units of token B. This generates a temporary price change from price impact.	$\alpha_t^A - \Delta_t^A$	$\alpha_t^B + \Delta_t^B$	✓	✓	✓
3	t	The yield farmer buys token B in a liquid pool for B using the rest of his/her fund.	$\alpha_t^A - \Delta_t^A$	$\alpha_t^B + \Delta_t^B$	✓	✓	
4	t	Arbitrageurs correct the price by supplying Δ_t^A of token A and receiving Δ_t^B of token B.	α_t^A	α_t^B			
5	t	The yield farmer provides liquidity to the pool and receives LP tokens. Denote the fraction of his/her tokens to the tokens in the current pool by $s(f)$.	$(1+s(f))\alpha_t^A$	$(1+s(f))\alpha_t^B$	✓		
6	t	The yield farmer stakes the LP tokens in a farm.	$(1+s(f))\alpha_t^A$	$(1+s(f))\alpha_t^B$	✓		
7	$t+h$	h days elapse.	$(1+s(f))\alpha_{t+h}^A$	$(1+s(f))\alpha_{t+h}^B$			
8	$t+h$	The yield farmer receives (harvests) realized farm yields in CAKE tokens.	$(1+s(f))\alpha_{t+h}^A$	$(1+s(f))\alpha_{t+h}^B$	✓		
9	$t+h$	The yield farmer withdraws his/her LP tokens.	$(1+s(f))\alpha_{t+h}^A$	$(1+s(f))\alpha_{t+h}^B$	✓		
10	$t+h$	The yield farmer sells their CAKE tokens.	$(1+s(f))\alpha_{t+h}^A$	$(1+s(f))\alpha_{t+h}^B$	✓	✓	
11	$t+h$	The yield farmer redeems their LP tokens at the liqudity pool and receives his/her shares of token A and B.	α_{t+h}^A	α_{t+h}^B	✓		
12	$t+h$	The yield farmer sells his/her $\Delta_{t+h}^A = s(f)\alpha_{t+h}^A$ of token A using the same pool. This generates a temporary price change from price impact. They receive Δ_{t+h}^B of token B in exchange from the liquidity pool.	$\alpha_{t+h}^A + \Delta_{t+h}^A$	$\alpha_{t+h}^B - \Delta_{t+h}^B$	✓	✓	✓
13	$t+h$	The yield farmer sell his/her $(\Delta_{t+h}^B + s(f)\alpha_{t+h}^B)$ of token B in a liquid pool for B.	$\alpha_{t+h}^A + \Delta_{t+h}^A$	$\alpha_{t+h}^B - \Delta_{t+h}^B$	✓	✓	
14	$t+h$	Arbitrageurs correct the price by supplying Δ_{t+h}^B of token B and receiving Δ_{t+h}^A of token A. A new round of yield farming starts again.	α_{t+h}^A	α_{t+h}^B			

Table A.2: Token Pairs used for Conversion of Token Prices into U.S. Dollars

In this table, we list tokens used as numeraires in the liquidity pools included in our data. *Frequency* refers to the number of liquidity pools in which the corresponding token is used as the numeraire. *PancakeSwap Token Pairs Used* lists the token pairs in PancakeSwap liquidity pools used for the conversion of token prices into U.S. dollars. *Data from CoinMarketCap* lists the tokens whose prices we collect from CoinMarketCap.

Numeraire Token	**Numeraire Token Symbol**	**Frequency**	**PancakeSwap Token Pairs Used**	**Data from CoinMarketCap**
Wrapped Binance Coin	BNB	146	BNB-BUSD	BUSD
Binance USD	BUSD	50	None	BUSD
TerraUSD	UST	6	None	UST
Binance-Peg Ethereum	ETH	4	ETH-BNB, BNB-BUSD	BUSD
PancakeSwap Token	CAKE	4	Cake-BNB, BNB-BUSD	BUSD
Binance-Peg Bitcoin	BTCB	3	BTCB-BNB, BNB-BUSD	BUSD
Binance-Peg Tether	USDT	2	None	USDT
Binance-Peg USD Coin	USDC	2	None	USDC
pTokens Bitcoin	PBTC	1	PBTC-BNB, BNB-BUSD	BUSD
QIAN Governance Token	KUN	1	KUN-BUSD	BUSD
Total		219		

Table A.3: Summary Statistics of Coins used for Constructing Cryptocurrency Factors

In this table, we provide summary statistics of cryptocurrencies used for construction of cryptocurrency factors. Our sample period for cryptocurrency factors starts on December 28, 2013 and ends on September 5, 2021. The unit for market capitalization and daily trading volume in this table is $ million.

Year	**# Coins**	**Market Capitalization**		**Daily Trading Volume**	
		Mean	**Median**	**Mean**	**Median**
2013	26	409.8	7.3	2.01	0.05
2014	100	260.1	4.1	1.21	0.03
2015	79	136.9	2.8	1.13	0.10
2016	157	171.5	3.5	1.76	0.02
2017	675	427.9	9.9	17.89	0.13
2018	1,250	415.8	10.9	23.64	0.15
2019	1,175	227.8	6.0	68.67	0.18
2020	1,520	301.0	6.8	121.25	0.29
2021	2,291	724.8	13.9	146.86	0.53

Table A.4: Comparison of Cryptocurrency Three-Factor Regressions

This table compares the regression results for portfolios sorted on one-week momentum by quintile. The sample period used in Liu, Tsyvinski, and Wu (2019) is from the beginning of 2014 to the end of 2018, which we interpret to be from the first week in 2014 to the 52nd (last) week of 2018.

Panel A: Regressions from Liu, Tsyvinski, and Wu (2019)			**Quintile**		
	1	**2**	**3**	**4**	**5**
α	-0.015	-0.010	-0.003	0.025	-0.012
$t(\alpha)$	*-1.970*	*-1.525*	*-0.657*	*1.470*	*-1.080*
β_{CMKT}	1.041	1.029	0.958	1.093	0.924
β_{CSMB}	0.124	0.014	0.204	0.072	0.297
β_{CMOM}	-0.164	-0.125	-0.071	0.072	0.424
R^2	0.531	0.606	0.687	0.198	0.435

Panel B: Replicated Regressions			**Quintile**		
	1	**2**	**3**	**4**	**5**
α	-0.019	-0.015	-0.004	0.031	-0.013
$t(\alpha)$	*-2.640*	*-2.362*	*-0.718*	*1.562*	*-1.230*
β_{CMKT}	0.994	0.957	0.873	1.119	0.996
β_{CSMB}	0.019	0.030	0.150	-0.034	0.081
β_{CMOM}	-0.148	-0.056	-0.045	-0.040	0.325
R^2	0.578	0.635	0.699	0.190	0.503

B.1 Appendix for Conceptual Framework

B.1 Capital gains and impermanent loss

In this section, we outline the procedure for deriving the main equations of Section 4 assuming that there are no trading frictions such as trading fees and gas fees. For expositional purposes, we consider the following scenario:

- Suppose a liquidity provider provides 1 BNB and 100 BUSD, a stablecoin pegged to U.S. dollar, to a liquidity pool.
- There is a total of 10 BNB and 1,000 BUSD in the pool after this liquidity provision. Therefore, the liquidity provider's share is 10%.
- After h days, the price of BNB becomes 200 BUSD.
- The liquidity provider withdraws his/her liquidity.

The constant product model imposes a condition that the product of two tokens should be constant. In this case, $k = \alpha_t^A \alpha_t^B = 10 \times 1,000 = 10,000$, where α^i is the number of cryptocurrency i in the liquidity pool. Let A and B be BNB and BUSD, repectively. Consider t as today and $t + h$ as h days after today. The value of A (BNB) in the pool should be identical to the value of B (BUSD) at any t, i.e. $P_t^A \alpha_t^A = P_t^B \alpha_t^B$ for all t. See Lemma 1 for more details.

Lemma 1) $P_t^A \alpha_t^A = P_t^B \alpha_t^B$ in a constant product model.

Proof) Under the constant product model, the product of the quantities of two cryptocurrencies should be constant, i.e. $\alpha_t^A \alpha_t^B = k$. This implies that $\frac{\partial \alpha_t^B}{\partial \alpha_t^A} = -\frac{\alpha_t^B}{\alpha_t^A}$. To purchase δ, a trader needs to pay $\delta \frac{\alpha_t^B}{\alpha_t^A}$., which means that $P_t^A \delta = P_t^B \delta \frac{\alpha_t^B}{\alpha_t^A} \rightarrow P_t^A \alpha_t^A = P_t^B \alpha_t^B$. $\square$

Since we have two equations: $P_t^A \alpha_t^A = P_t^B \alpha_t^B$ and $k = \alpha_t^A \alpha_t^B$, we can solve for both α_t^A and α_t^B:

$$\alpha_t^A = \sqrt{k \left(\frac{P_t^B}{P_t^A} \right)}, \alpha_t^B = \sqrt{k \left(\frac{P_t^A}{P_t^B} \right)}.$$

Given that the rate of exchange for 1 BNB becomes 200 BUSD (which is equivalent to \$200, assuming that BUSD is perfectly pegged to the U.S. dollar) at time $t+h$:

$$\alpha^A_{t+h} = \sqrt{k\left(\frac{P^B_{t+h}}{P^A_{t+h}}\right)} = \sqrt{10,000 \times (\$1/\$200)} = \sqrt{50} = 7.07,$$
$$\alpha^B_{t+h} = \sqrt{k\left(\frac{P^A_{t+h}}{P^B_{t+h}}\right)} = \sqrt{10,000 \times (\$200/\$1)} = \sqrt{2,000,000} = 1414.21.$$

The liquidity provider's share is 10%. If he/she withdraws their liquidity, he/she will get 0.707 BNB and 141.421 BUSD. This amounts to $0.707 \times 200 + 141.421 \times 1 = \$282.82$. In the crypto community, the impermanent loss is often defined as the percentage of the ratio of investment outcomes at time $t+h$ in two scenarios: (1) providing liquidity to the pool at t or (2) directly holding the underlying assets. If the liquidity provider simply held the assets (1 BNB and 100 BUSD), he/she would now have $\$300 = 1 \times 200 + 100 \times 1$ worth of assets. In this case, the impermanent loss is $(282.82/300 - 1) \times 100 =$ -5.727%. To formalize this, we compute the following measure which is the ratio of investment outcomes at time $t+h$ in two scenarios minus 1. In this example, the liquidity provider's share is 10%. Let's assume that his/her share in general is ω.

$$\frac{\omega(P^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h})}{\omega(P^A_{t+h}\alpha^A_t + P^B_{t+h}\alpha^B_t)} - 1$$

Note that α^i in the denominator is the same as the number of shares the liquidity provider initially held, whereas α^i in the numerator is the number of shares after trading activities between t and $t+h$.

$$\begin{aligned}
&\frac{\omega(P^A_{t+h}\alpha^A_{t+h}+P^B_{t+h}\alpha^B_{t+h})}{\omega(P^A_{t+h}\alpha^A_t+P^B_{t+h}\alpha^B_t)}-1\\
&=\frac{\left(\frac{P^A_{t+h}}{P^B_{t+h}}\right)\alpha^A_{t+h}+\alpha^B_{t+h}}{\left(\frac{P^A_{t+h}}{P^B_{t+h}}\right)\alpha^A_t+\alpha^B_t}-1\\
&=\frac{\left(\frac{P^A_{t+h}}{P^B_{t+h}}\right)\sqrt{k\left(\frac{P^B_{t+h}}{P^A_{t+h}}\right)}+\sqrt{k\left(\frac{P^A_{t+h}}{P^B_{t+h}}\right)}}{\left(\frac{P^A_{t+h}}{P^B_{t+h}}\right)\sqrt{k\left(\frac{P^B_t}{P^A_t}\right)}+\sqrt{k\left(\frac{P^A_t}{P^B_t}\right)}}-1\\
&=\frac{\left(\frac{P^A_{t+h}}{P^B_{t+h}}\right)\sqrt{\frac{P^B_{t+h}}{P^A_{t+h}}}+\sqrt{\frac{P^A_{t+h}}{P^B_{t+h}}}}{\left(\frac{P^A_{t+h}}{P^B_{t+h}}\right)\sqrt{\frac{P^B_t}{P^A_t}}+\sqrt{\frac{P^A_t}{P^B_t}}}-1
\end{aligned}$$

Denote the relative price of token A to token B at t by ρ_t $(=\frac{P^A_t}{P^B_t})$. Then, the above expression is reduced to

$$\frac{\rho_{t+h}\sqrt{\frac{1}{\rho_{t+h}}}+\sqrt{\rho_{t+h}}}{\rho_{t+h}\sqrt{\frac{1}{\rho_t}}+\sqrt{\rho_t}}-1=\frac{2\sqrt{\rho_{t+h}}}{\rho_{t+h}\sqrt{\frac{1}{\rho_t}}+\sqrt{\rho_t}}-1=\frac{2\sqrt{\frac{\rho_{t+h}}{\rho_t}}}{\frac{\rho_{t+h}}{\rho_t}+1}-1.$$

The above figure shows the relation between change of the relative price $(\frac{\rho_{t+h}}{\rho_t})$ and the impermanent loss, defined as the ratio of investment outcomes at time $t+h$ in the two scenarios, minus 1. If ρ changes and $\frac{\rho_{t+h}}{\rho_t}$ deviates from 1, the liquidity provider experiences a loss compared to a simple position of holding underlying tokens from t. It is straightforward to show that this loss, also called the impermanent loss, is non-positive: $\frac{2\sqrt{\frac{\rho_{t+h}}{\rho_t}}}{\frac{\rho_{t+h}}{\rho_t}+1}-1=-\frac{(\sqrt{\frac{\rho_{t+h}}{\rho_t}}-1)^2}{\frac{\rho_{t+h}}{\rho_t}+1}$.

However, the above approach is not directly applicable to our analysis because we analyze returns from liquidity provision, rather than comparing an investment outcome at $t+h$ with an investment outcome in a hypothetical situation at $t+h$. Therefore, our goal is to simplify the liquidity provider's gross return, expressed as follows:

$$\frac{\omega(P^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h})}{\omega(P^A_t\alpha^A_t + P^B_t\alpha^B_t)}$$

We can decompose the above expression into two parts.

$$\begin{aligned}
&\frac{P^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h}}{P^A_t\alpha^A_t + P^B_t\alpha^B_t} \\
&= \left(\left(\frac{P^A_t\alpha^A_t}{P^A_t\alpha^A_t + P^B_t\alpha^B_t}\right)R^A_{t,t+h} + \left(\frac{P^B_t\alpha^B_t}{P^B_t\alpha^A_t + P^B_t\alpha^B_t}\right)R^B_{t,t+h}\right) \\
&+ \left(\frac{P^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h}}{P^A_t\alpha^A_t + P^B_t\alpha^B_t} - \left(\left(\frac{P^A_t\alpha^A_t}{P^A_t\alpha^A_t + P^B_t\alpha^B_t}\right)R^A_{t,t+h} + \left(\frac{P^B_t\alpha^B_t}{P^B_t\alpha^A_t + P^B_t\alpha^B_t}\right)R^B_{t,t+h}\right)\right)
\end{aligned}$$

We call the first term capital gains, which is a return that an investor can earn if he/she holds α^A_t and α^B_t shares of token A and B until time $t+h$ without providing liquidity to the pool. We define the second term as impermanent loss in our context, which is the difference between the return on liquidity provision and capital gains.

First, the capital gains are reduced to $\frac{1}{2}R^A_{t,t+h} + \frac{1}{2}R^B_{t,t+h}$ thanks to Lemma 1. Second, in order to simplify the impermanent loss, we use Lemma 1 again.

$$\begin{aligned}
&\frac{P^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h}}{P^A_t\alpha^A_t + P^B_t\alpha^B_t} - \left(\left(\frac{P^A_t\alpha^A_t}{P^A_t\alpha^A_t + P^B_t\alpha^B_t}\right)R^A_{t,t+h} + \left(\frac{P^B_t\alpha^B_t}{P^B_t\alpha^A_t + P^B_t\alpha^B_t}\right)R^B_{t,t+h}\right) \\
&= \frac{P^A_{t+h}\alpha^A_{t+h}}{P^A_t\alpha^A_t} - \left(\frac{1}{2}R^A_{t,t+h} + \frac{1}{2}R^B_{t,t+h}\right) \\
&= \frac{P^A_{t+h}\sqrt{k\left(\frac{P^B_{t+h}}{P^A_{t+h}}\right)}}{P^A_t\sqrt{k\left(\frac{P^B_t}{P^A_t}\right)}} - \left(\frac{1}{2}R^A_{t,t+h} + \frac{1}{2}R^B_{t,t+h}\right) \\
&= \sqrt{R^A_{t,t+h}R^B_{t,t+h}} - \left(\frac{1}{2}R^A_{t,t+h} + \frac{1}{2}R^B_{t,t+h}\right) \\
&= -\frac{1}{2}\left(\sqrt{R^A_{t,t+h}} - \sqrt{R^B_{t,t+h}}\right)^2
\end{aligned}$$

The impermanent loss defined in the context of return on liquidity provision is closely related to the impermanent loss defined as the ratio of investment outcomes at time $t+h$

in the two scenarios minus 1. It is straightforward to show that

$$-\frac{1}{2}\left(\sqrt{R^A_{t,t+h}}-\sqrt{R^B_{t,t+h}}\right)^2=\left(\frac{1}{2}R^A_{t,t+h}+\frac{1}{2}R^B_{t,t+h}\right)\left(\frac{2\sqrt{\frac{\rho_{t+h}}{\rho_t}}}{\frac{\rho_{t+h}}{\rho_t}+1}-1\right)$$

B.2 Trading frictions in yield farming

In this section, we explain trading frictions in yield farming and how they can affect the performance of a yield farming strategy. We investigate three different trading frictions: gas fees, trading fees, and price impact.

Gas fees

Table A.1 lists 14 steps for one round of the yield farming strategy. Out of the 14 steps, 10 require the farmer to pay gas fees. The gas fee is the transaction cost that BSC users need to pay whenever they execute transactions that require computational resources of the network. The gas fee is typically not proportional to the size of the transaction. We discuss in Section 5.2 about how we collect gas fee data in yield farming. We subtract the gas fee in each round of yield farming from invested capital to incorporate the effect of gas fee on the performance of yield farming.

Trading fees

Let c^*=0.0025 (0.25%) denote the fraction of trading volume that traders need to pay in trading fees at PancakeSwap. In Step 2, when a yield farmer buys token A, he/she pays a 0.25% trading fee. Because this trading fee does not apply to token B, the farmer pays effectively half of the trading fee $\frac{c*}{2}$ (=0.125%) of the additional liquidity that he/she provides. Moreover, the farmer has to pay an additional fee of 0.25%: when he/she converts the withdrawn token A to token B, he/she pays additional $\frac{c*}{2}$ fraction of trading fee. The farmer also needs to pay also $\frac{c*}{2}$ of trading fee in Step 3 and Step 13. Given that the yield farmer pays $\frac{c*}{2}$ our times, the yield farmer's gross return on capital gain and impermanent loss should be

$$\frac{(1-2c^*)\left(P^A_{t+h}\alpha^A_{t+h}+P^B_{t+h}\alpha^B_{t+h}\right)}{P^A_t\alpha^A_t+P^B_t\alpha^B_t}=(1-2c^*)\left(\left(\frac{1}{2}R^A_{t,t+h}+\frac{1}{2}R^B_{t,t+h}\right)-\frac{1}{2}\left(\sqrt{R^A_{t,t+h}}-\sqrt{R^B_{t,t+h}}\right)^2\right).$$

In Step 10, the yield farmer also needs to pay trading fees when he/she sells CAKE tokens harvested from farming. For this, we multiply $(1 - c^*)$ on the realized farm yield term in equation (10).

Price impact

Executing a yield farming strategy involves buying and selling token A. In Step 2, a yield farmer buys token A. As a result of price impact, the yield farmer will buy token A at a price above the current market price. Symmetrically, the yield farmer will sell token A at a price below the current market price. Such adverse price impacts will result in additional losses for the yield farmer. The size of the loss is proportional to the relative size of investment (I_t) to the size of the liquidity pool, i.e., $I_t = fL_t$. We go through each step in Table A.1 to investigate the price impacts involved in a yield farming strategy.

(1) Step 1: We start from a liquidity pool with two tokens A and B. It has α_t^A of token A and α_t^B of token B and the prices of token A and B are denoted as P_t^A and P_t^B.

(2) Step 2: An yield farmer buys Δ_t^A number of token A using a part of his/her fund, $I_t = fL_t$. What is important here is that the yield farmer must obtain tokens A and B proportionally to α_t^A/α_t^B. For this purpose, we divide his/her fund into xI_t and $(1-x)\,I_t$ to allocate towards token A and B, respectively. The yield farmer first converts $\$xI_t$ to token B in a liquid market for B. Then, the farmer will have $\frac{xI_t}{P_t^B}$ of token B on hand, which he/she will convert to token A by means of the liquidity pool. Due to the constant product model assumption,

$$\left(\alpha_t^A - \Delta_t^A\right)\left(\alpha_t^B + \frac{xI_t}{P_t^B}\right) = \alpha_t^A \alpha_t^B$$

If we solve this for Δ_t^A,

$$\Delta_t^A = \frac{\left(\frac{xI_t}{P_t^B}\right)\alpha_t^A}{\alpha_t^B + \frac{xI_t}{P_t^B}} = \frac{xI_t\alpha_t^A}{P_t^B\alpha_t^B + xI_t} = \frac{xI_t\alpha_t^A}{\frac{1}{2}L_t + xI_t} = \frac{xf\alpha_t^A}{\frac{1}{2} + xf}$$

(3) Step 3: The yield farmer uses the rest of their funds, $\$(1-x)\,I_t$, to buy token B in a liquid market for B. Then, he/she will get Δ_t^B of token B where Δ_t^B is expressed as follows.

$$\Delta_t^B = \frac{(1-x)\,I_t}{P_t^B} = \frac{(1-x)\,fL_t}{P_t^B}.$$

Finally, we find x that satisfies $\frac{\Delta_t^A}{\Delta_t^B} = \frac{\alpha_t^A}{\alpha_t^B}$.

$$\frac{\Delta_t^A}{\Delta_t^B} = \frac{\frac{xf\alpha_t^A}{\frac{1}{2}+xf}}{\frac{(1-x)fL_t}{P_t^B}} = \frac{\frac{xf\alpha_t^A}{\frac{1}{2}+xf}}{\frac{(1-x)f\left(2P_t^B\alpha_t^B\right)}{P_t^B}} = \left(\frac{x}{1-x}\right) \left(\frac{1}{1+2xf}\right) \frac{\alpha_t^A}{\alpha_t^B}.$$

Therefore,

$$\left(\frac{x}{1-x}\right) \left(\frac{1}{1+2xf}\right) = 1.$$

If we solve for x,

$$x = \frac{f-1+\sqrt{f^2+1}}{2f}.$$

There are two solutions, but only the above solution is positive.

(4) Step 4: Arbitrageurs correct the price by supplying Δ_t^A of token A and receiving Δ_t^B of token B in return, after which the liquidity pool becomes basically identical to the initial pool.

(5) Step 5: The yield farmer provides their liquidity to the pool and receives LP tokens. For simplicity of notation, let's define $s(f)$, the ratio of the yield farmer's share to the current share in the liquidity pool before the yield farmer provides the liquidity.

$$s(f) = \frac{\Delta_t^A}{\alpha_t^A} = \frac{\frac{xI_t\alpha_t^A}{\frac{1}{2}L_t+xI_t}}{\alpha_t^A} = \frac{xfL_t}{\frac{1}{2}L_t+xfL_t} = \frac{f \times \left(\frac{f-1+\sqrt{f^2+1}}{2f}\right)}{\frac{1}{2}+f\times\frac{f-1+\sqrt{f^2+1}}{2f}} = \frac{f-1+\sqrt{f^2+1}}{f+\sqrt{f^2+1}}$$

After the liquidity provision by the yield farmer, the shares of token A and B become $\alpha_t^A(1+s(f))$ and $\alpha_t^B(1+s(f))$. Now, we measure the price impact when the yield farmer buys Δ_t^A of token A. The farmer uses $\$xI_t$ to buy Δ_t^A of token A. This means that the effective price paid by the farmer is:

$$\tilde{P}_t^A = \frac{xI_t}{\Delta_t^A} = \frac{xfL_t}{\frac{xf\alpha_t^A}{\frac{1}{2}+xf}} = \frac{xf\left(2P_t^A\alpha_t^A\right)}{\frac{xf\alpha_t^A}{\frac{1}{2}+xf}} = 2P_t^A\left(\frac{1}{2}+xf\right) = P_t^A(1+2fx)$$

$$= P_t^A\left(1+\left(f-1+\sqrt{f^2+1}\right)\right)$$

Given that $f-1+\sqrt{f^2+1} > 0$, $\tilde{P}_t^A > P_t^A$.

(6) Step 6: The yield farmer stakes the P tokens to a farm.

(7) Step 7: The yield farmer waits for h days. After the trading activities the shares of token A and B become $\alpha_{t+h}^A(1+s(f))$ and $\alpha_{t+h}^B(1+s(f))$.

(8) Step 8: The yield farmer receives (harvest) realized farm yields in Cake tokens.

(9) Step 9: The yield farmer withdraws his/her LP tokens from the farm.

(10) Step 10: The yield farmer sells Cake tokens.

(11) Step 11: The yield farmer withdraws his/her liquidity from the liquidity pool by sending the LP tokens to the pool. After the farmer's withdrawing liquidity, the shares of token A and B in the pool become α_{t+h}^A and α_{t+h}^B.

(12) Step 12: The yield farmer sells his $\Delta_{t+h}^A = s(f)\alpha_{t+h}^A$ of token A and receives Δ_{t+h}^B of token B. Currently, there are α_{t+h}^A and $\alpha_{t+h}^B$of token A and token B in the pool. After the farmer's sending $\Delta_{t+h}^A = s(f)\alpha_{t+h}^A$ of token A, he/she receives Δ_{t+h}^B token B. Due to the constant product model,

$$\left(\alpha_{t+h}^A + s(f)\alpha_{t+h}^A\right)\left(\alpha_{t+h}^B - \Delta_{t+h}^B\right) = \alpha_{t+h}^A\alpha_{t+h}^B$$

$$\rightarrow \Delta_{t+h}^B = \frac{s(f)}{1+s(f)}\alpha_{t+h}^B$$

The farmer sends $s(f)\alpha_{t+h}^A$ of token A and in return $P_{t+h}^B\Delta_{t+h}^B$ worth of USD. This means that the effective price that the yield farmer receives when selling token A is

$$\tilde{P}_{t+h}^A = \frac{P_{t+h}^B\Delta_{t+h}^B}{s(f)\alpha_{t+h}^A} = \frac{\frac{s(f)}{1+s(f)}\alpha_{t+h}^B P_{t+h}^B}{s(f)\alpha_{t+h}^A} = \frac{\frac{s(f)}{1+s(f)}\alpha_{t+h}^A P_{t+h}^A}{s(f)\alpha_{t+h}^A} = \left(\frac{1}{1+s(f)}\right)P_{t+h}^A$$

So the yield farmer sells at a lower price than P_{t+h}^A.

(13) Step 13: The yield farmer sells $\Delta_{t+h}^B + s(f)\alpha_{t+h}^B$ of token B in a liquid market for token B.

(14) Step 14: An arbitrageur corrects the price by supplying Δ_{t+h}^B of token B and receiving Δ_{t+h}^A of token A. A new round of yield farming starts again.

Now we compute the return of this yield farming strategy considering the price impact. First, the yield farmer uses his/her fund $I_t = fL_t = \tilde{P}_t^A\left(s(f)\alpha_t^A\right) + P_t^B(s(f)\alpha_t^B)$ to buy $s(f)\alpha_t^A$ and $s(f)\alpha_t^B$ shares of token A and B at $\tilde{P}_t^A$ and P_t^B. After h days, the yield farmer withdraws $s(f)\alpha^A_{t+h}$ and $s(f)\alpha^B_{t+h}$ shares of token A and B and sell them at $\tilde{P}^A_{t+h}$ and P^B_{t+h}. Then, its gross return is expressed as

$$\frac{\tilde{P}^A_{t+h}\left(s(f)\alpha^A_{t+h}\right) + P^B_{t+h}(s(f)\alpha^B_{t+h})}{\tilde{P}^A_t\left(s(f)\alpha^A_t\right) + P^B_t(s(f)\alpha^B_t)} = \frac{\tilde{P}^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h}}{\tilde{P}^A_t\alpha^A_t + P^B_t\alpha^B_t}.$$

We simplify this as follows.

$$\begin{aligned}
\frac{\tilde{P}^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h}}{\tilde{P}^A_t\alpha^A_t + P^B_t\alpha^B_t} =& \frac{\left(\frac{1}{1+s(f)}\right)P^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h}}{P^A_t\left(1+\left(f-1+\sqrt{f^2+1}\right)\right)\alpha^A_t + P^B_t\alpha^B_t} \\
=& \frac{\left(\frac{1}{1+s(f)}+1\right)P^A_{t+h}\alpha^A_{t+h}}{\left(1+\left(f-1+\sqrt{f^2+1}\right)+1\right)P^A_t\alpha^A_t} \\
=& \frac{\frac{1}{1+s(f)}+1}{f+1+\sqrt{f^2+1}}\left(\frac{P^A_{t+h}\alpha^A_{t+h}}{P^A_t\alpha^A_t}\right) \\
=& \lambda(f)\left(\frac{P^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h}}{P^A_t\alpha^A_t + P^B_t\alpha^B_t}\right) \\
=& \lambda(f)\left(\left(\frac{1}{2}R^A_{t+h} + \frac{1}{2}R^B_{t+h}\right) - \frac{1}{2}\left(\sqrt{R^A_{t+h}} - \sqrt{R^B_{t+h}}\right)^2\right)
\end{aligned},$$

where

$$\lambda(f) = \frac{\frac{1}{1+s(f)}+1}{f+1+\sqrt{f^2+1}} = \frac{\frac{1}{1+\frac{f-1+\sqrt{f^2+1}}{f+\sqrt{f^2+1}}}+1}{f+1+\sqrt{f^2+1}} = \frac{3f+3\sqrt{f^2+1}-1}{\left(2f+2\sqrt{f^2+1}-1\right)\left(f+1+\sqrt{f^2+1}\right)},$$

In sum, if we take into account of both price impact and trading fee, return will be

$$(1-2c^*)\lambda(f)\left(\left(\frac{1}{2}R^A_{t+h} + \frac{1}{2}R^B_{t+h}\right) - \frac{1}{2}\left(\sqrt{R^A_{t+h}} - \sqrt{R^B_{t+h}}\right)^2\right)$$

where $(1-2c^*)\lambda(f) < 1$.

Figure 3 illustrates the price impact in buying and selling token A and $\lambda(f)$, which summarizes the overall effect of price impacts on the performance of yield farming. Panel A shows the relation between f and $\frac{\tilde{P}^A_t}{P^A_t}$. $\frac{\tilde{P}^A_t}{P^A_t}$ is greater than or equal to 1 and increasing in f, which implies that the yield farmer pays higher prices than the current market price when they purchase token A, which is attenuated as the size of his/her investment increases. Panel B shows the relationship between f and $\frac{\tilde{P}^A_{t+h}}{P^A_{t+h}}$. This is less than or equal to 1 and decreasing in f, which means that the yield farmer sells token A at a larger discount as the size of investment increase. Finally, Panel C plots $\lambda(f)$ with respect to f. $\lambda(f)$ is less than or equal to 1, decreasing in f, and its effect is substantial when f is large. For example, if the yield farmer's investment is very small such that f is close 0, $\lambda(f) = 1$ and therefore, there is no effect. However, if the yield farmer invests as much as the size of the pool ($f = 1$), he/she will lose more than 50% of their gross return.

C.1 Accuracy of Constructed Cryptocurrency Factors

As a test of the accuracy of our methodology, we replicate the three-factor regressions from Table 11 in Liu, Tsyvinski, and Wu (2019) for portfolios sorted on one-week momentum by quintile, a set of implementable trading strategies not used in the construction of the three factors. In table A.4, we compare our parameter estimates to those obtained in Liu, Tsyvinski, and Wu (2019). The two are near-identical with only minor deviations, which may be from small variations in the sample period used and/or changes in the markets for which Coinmarketcap tracks price data.

In addition, it is worth noting that the estimates for alpha obtained in Liu, Tsyvinski, and Wu (2019) are reported in weekly frequency, whereas our measures of alpha have been annualized. For instance, a weekly alpha of 0.025, as is the case for the fourth quintile of one-week momentum in table A.4, translates into a yearly alpha of 2.611 when annualized. Therefore, the magnitudes of our estimates of alpha for yield-farming strategies are reasonably comparable to strategies analyzed in table 11 of Liu, Tsyvinski, and Wu (2019), in which three-factor weekly alphas exceed 0.02 (or an annualized alpha of 1.80) for many price- and momentum-based strategies.