Reaching for Yield in Decentralized Financial Markets[*]

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

First draft: October 15, 2021
This draft: March 21, 2024

Abstract

Yield farms in decentralized finance dynamically compete for liquidity by offering high yields, advertised as salient headline rates. Farming these yields involves complex investment strategies with hidden downside risks. Capitalizing on the transparency of blockchain transactions data, we show that investors chase farms with high yields and that farms with the highest headline rates record the most negative risk-adjusted returns. Through randomized shocks to yield farmers' information display, we show that improved risk disclosure and lower perceived product complexity reduces yield chasing, thereby improving investor performance. Our evidence is consistent with salience theory that may underpin reaching for yield behavior.

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

Keywords: complexity, decentralized finance, derivatives, reaching for yield, salience

[*]As for helpful comments, we thank Matteo Aquilina (discussant), Greg Brown, Mikhail Chernov, Kyoung Jin Choi (discussant), Kim-Sau Chung, William Cong, Darrell Duffie, Eric Ghysels, Campbell Harvey, Felix Irresberger (discussant), Douglas Laporte (discussant), Valerie Laturnus (discussant), Alfred Lehar, Ljubica Georgievska (discussant), Shimon Kogan (discussant) Christian Lundblad, Igor Makarov, Paige Ouimet, Neil Pearson, Cameron Peng (discussant), Daniel Rabetti (discussant), Adam Reed, Alessio Saretto, Lorenzo Schönleber (discussant), conference 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 Conference, the Second Annual DeFi Academic Research Conference (organized by Duke and UNC), the 16th Financial Risks International Forum on "Finance & Society," 2023 Financial Intermediation Research Society Conference, the 9th BI-SHoF Conference on Asset Pricing and Financial Econometrics, the 2023 Northern Finance Association Annual Meeting, the 2023 WBS Gillmore Centre Conference on DeFi & Digital Currencies, the 2023 EUROFIDAI-ESSEC Paris December Finance Meeting, the 2nd Bonn/Mannheim workshop on digital finance, the UD/Philly Fed Fintech & Financial Institutions Conference, and seminar participants at the Bank of Canada, McGill University, UNC Kenan-Flagler Business School, Joint Digital-Economy Seminar by Monash Univ., HKBU, and National Taiwan Univ., Renmin Univ., and KAIST/Korea University Business School, the Frankfurt School of Finance & Management, ESSEC Business School, CBER Symposium - Investment in DeFi, York University, HEC Paris, University of Washington Foster School of Business, EDHEC Business School, Johns Hopkins University Carey School of Business, NOVA SBE, Warwick Business School, INSEAD, Tilburg University, Maastricht University and the University of St. Gallen. We also thank Wei Yang for excellent research assistance. Augustin acknowledges financial support from the Canadian Derivatives Institute and from the Canada Research Chair Program of the Social Sciences and Humanities Research Council Canada. Shin acknowledges financial support from the Kenan Institute. All authors acknowledge support from the Chaire Fintech AMF-Finance Montréal: "*Ce projet a été réalisé dans le cadre des activités de recherche de la Chaire Fintech AMF-Finance Montréal dont le soutien financier provient de l'Autorité des marchés financiers et de Finance Montréal. Les informations, opinions et avis exprimés n'engagent que la responsabilité de leurs auteurs.*" *A prior draft of this paper was circulated under the title 'Yield Farming'.*

[†]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.

"Crypto 'yield farmers' chase high returns, but risk losing it all."
Alexander Osipovich, Wall Street Journal

"We just don't have enough investor protection in crypto [...], 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 (Harvey, Ramachandran, and Santoro, 2021; Makarov and Schoar, 2022; John, Kogan, and Saleh, 2023). Operating through applications built on blockchains and executed through smart contracts, DeFi intends to counteract the influence of traditional centralized financial intermediaries.

Figure 1 illustrates that total value locked (TVL) in DeFi, a measure of aggregate capital invested in decentralized financial applications, grew exponentially to almost \$200 billion in less than 2 years. Despite the sharp drop associated with a general devaluation of digital currencies in the summer of 2022, Figure 1 shows that the number of active applications with TVL above \$1 million has remained high, close to 700 DeFi platforms.

The rapid growth of DeFi has raised regulatory concerns. One concern originates from DeFi platforms competing for liquidity provision through offering extraordinarily high yields while exposing investors to significant downside risks (e.g., Oliver, 2021; Osipovich, 2021; Kruppa, 2022). Moreover, DeFi platforms bear resemblance to complex structured retail products and are easily accessible to retail investors despite their product complexity. The Securities and Exchange Commission refers to certain investments as 'unregulated and complex strategies', with 'hidden risks to unsophisticated investors' (e.g., Gensler, 2021).

In this paper, we study yield farming, a decentralized financial service that is well-suited for examining investor behavior in the presence of product complexity. First, yield farms dynamically compete for liquidity provision by offering high yields to investors. These yields are salient and aggressively marketed as headline rates without disclosure of transaction costs, past performance, or potential downside risks. Second, yield farming is complex in both execution and payoffs, with hidden risks that are not well understood, according to survey evidence. Finally, we observe the entire history of transactions from blockchain data and can dynamically study investor behavior, including investment size, mistakes, and their response to changes in information disclosure and perceived product complexity.

Our overall evidence is supportive of the key features of salience theory (e.g., Bordalo, Gennaioli, and Shleifer, 2012, 2013, 2016, 2022). Yield farms promise passive income at impressive headline rates and investors chase farms with high yields. High yield farms also appear to have shrouded risk attributes (Gabaix and Laibson, 2006), since farms with the highest promised yields record the worst risk-adjusted performance ex-post. We find that this underperformance is amplified for small investment stakes and investor mistakes.

We first provide a conceptual framework for understanding the risk-return trade-offs of yield farming. Yield farming is a mechanism for passively earning income by supplying digital liquidity. While farming looks simple and accessible, with salient high yields, it involves a long chain of interlinked transactions subject to complexity in both execution and payoffs.

To become yield farmers, investors first need to act as digital liquidity providers. That requires the provision of pairs of cryptocurrency tokens in equal dollar amounts to a liquidity pool. Investors can choose among a menu of liquidity pools, each one associated with a pair of cryptocurrency tokens. The liquidity provision is certified through a liquidity token that represents the fractional ownership of the aggregate liquidity in the pool.

Investors can increase their passive earnings by staking the liquidity token to a yield farm. Each liquidity pool is linked to a unique farm that promises a salient interest rate often exceeding one hundred percent. That yield, which is paid using the governance token of the yield farming platform, is a complex function of farm and aggregate market characteristics. Paradoxically, the governance token owners maintain centralized voting power to adjust each farm's yield multiplier, which is one salient component of the yield function that can be influenced to dynamically compete for liquidity.

Yield farming performance can be decomposed into four components. First, the initial liquidity provision is rewarded through trading fees collected from third party traders buying and selling cryptocurrency tokens in a liquidity pool. Second, investors are exposed to the buy-and-hold price risk of the pledged tokens. Third, liquidity miners face significant downside risk through impermanent losses, which are defined through a loss function that non-linearly depends on the return correlation of the cryptocurrency pair. Fourth, yield farmers earn passive income in proportion to the aggregate liquidity locked in a yield farm.

Three types of transaction costs significantly alter yield farming performance. Each transaction requires the payment of a flat gas fee, implying that small investments are penalized by large overhead costs. Second, large investments relative to the existing liquidity result in significant price impact, especially at redemption. These observations suggest the existence of a trade-off that involves an optimal investment size. Finally, since it is strictly dominating to fully pledge the liquidity tokens into yield farms, staking ratios below one reduce investment performance and are a sign of investor mistakes.

In a second step, we provide new stylized facts on yield farms, investor behavior, and investment performance. Our analysis is based on a novel hand-collected data set of 262 yield farms from PancakeSwap, a yield farm platform hosted on the Binance Smart Chain (BSC), between March 1, 2021 and July 31, 2022. We focus on PancakeSwap because it is the largest yield farm ecosystem, with 435,130 active users on October 24, 2021, compared to 47,730 active users recorded on Uniswap. In addition, BSC features high trade execution speeds, lower congestion risks and lower trading fees than other comparable blockchains like Ethereum, making it more easily accessible to retail investors. Figure 2 indeed illustrates that gas fees paid for blockchain transactions are an order of magnitude larger for Ethereum.

There is significant heterogeneity in offered yields among the 262 farms in our sample. The average (median) offered yield is 80.53% (47.43%) with a standard deviation of 85.54%. These yields are salient and advertised as headline rates in enticing ways that feature cartoons, rockets, or emojis. In contrast, information on past performance and impermanent losses is hidden and challenging to find. Investing into yield farms is complex both in payoff and complexity. There are three underlying assets, non-linearities, and a full round-trip cost can take up to 14 transactions.

Offered farm yields are driven by five components related to the issuance of the platform's governance token CAKE, its price, which is common across all farms, each farm's liquidity, a farm multiplier, and an aggregate farm multiplier. Governance token owners may vote to increase or decrease farm multipliers as an instrument to incentivize liquidity provision. We find that the component of yield changes associated with multiplier changes is positively related to past trading fees and negatively to past realized yields. In addition, we observe that farms are delisted in response to low liquidity and weak trading fee revenue.

The examination of transaction records on the blockchain suggests that many yield farmers are financially unsophisticated. First, we observe that many investors do not migrate their funds when PancakeSwap switched to a newer and more secure platform in April 2021, even though the new platform would mechanically provide superior return potential. We see similar patterns when PancakeSwap migrated its staking functionality to a new staking contract in April 2022. Second, in spite of an optimal yield farm staking ratio of one, we find that the median staking ratio is below one most of the time.

The farmer data further suggest that the average yield farmer invests about $7,732 in 2.64 farms. Strikingly, we observe that smaller investment stakes are correlated with smaller staking ratios, suggesting that retail investors are more likely to leave money on the table. Survey evidence of 1,347 yield farmers also suggests that many investors lack financial sophistication, since 79% of them claim to understand the associated risks and rewards of yield farming, while only 33% state that they understand impermanent loss.

We next assess the return performance of yield farming strategies. Without transaction costs, yield farming appears to be profitable on average, with Sharpe ratios that are similar (but higher) to those of investments into the S&P500 index or Ethereum. Sorting farms into quintiles based on the magnitude of the offered yield reveals that high yield farms systematically generate the lowest returns because they incur the greatest impermanent loss, which is the hidden downside risk that is poorly understood. We further show that farmers who invest in higher yielding farms underperform by an additional 23bps for every 100% increase in offered yields ($\approx$ one standard deviation). Our overall evidence suggests that farms with the highest headline rates exhibit the worst risk-adjusted performance.

High yield farms are also those where investor mistakes have the most severe consequences since more money is left on the table in the absence of yield farming using the LP tokens. Accounting for transaction costs such as gas fees, trading fees and price impact further reduces the return performance across all yield quintiles.

Third, we provide evidence that investors exhibit yield chasing behavior that can result in negative risk-adjusted returns. Specifically, we identify all cases where Pancake token owners vote on changing the yield multiplier of one farm without significant changes to the multipliers of competing farms. In a difference-in-differences setting, we show that the differential increase (decrease) in aggregate farm flows in response to multiplier increases (decreases) is about 18%-19% (10%-13%), depending on the measurement of flows. A systematic analysis on the relation between flows and farm yields suggests that high headline yields predict positive net inflows, while flows are insensitive to impermanent losses.

At the farmer level, we document a positive propensity to buy riskier assets. We find that the average farmer provides about 2.55 percentage points more liquidity to a farm if it offers a 100% larger yield. Because high yield farms exhibit the worst risk-adjusted returns, our evidence is consistent with reaching for yield behavior. We also find that experience, as measured by the number of investment farms and the farming duration, reduces the reaching for yield propensity by 20% to 39%.

As a last step, we capitalize on a unique setting in PancakeSwap to study the impact of information disclosure and perceived complexity reduction on reaching for yield behavior. Yieldwatch, a third-party information platform, summarizes statistics on investor performance, such as historical capital gains and impermanent losses of individual farmers, and discloses it conditional on the acquisition of Yieldwatch tokens. Using the comprehensive trading history of individual investors, including their acquisitions of Yieldwatch tokens, we show that the enhanced information disclosure and reduction in complexity alleviates the intensity of yield-chasing behavior by about 58%, thereby improving the overall investor performance. This effect is also present in a quasi-natural experiment which exploits the random token allocation to successful and unsuccessful bidders around an unpredictable bidding threshold in the Yieldwatch initial token offering.

We confirm these findings in a different setting using airdrops organized by APY.Vision, which provides similar functions to Yieldwatch, but randomizes the acquisition of tokens needed to access the information platform. This analysis is implemented on a different yield farming platform, SushiSwap, built on the Ethereum blockchain, and, therefore, supports external validity of our findings. Overall, this evidence has important implications for information disclosure and investor protection in markets for high-yielding financial securities.

2 Literature

Our work relates to theories on financial innovation and security design. One view is that financial securities can be tailored to complete the market and, therefore, improve risk sharing (Allen and Gale, 1994; Duffie and Huang, 1995). Another view is that, when investors have salient preferences (Bordalo, Gennaioli, and Shleifer, 2012, 2013, 2022), financial intermediaries may compete by attracting consumers based on salient price attributes. An equilibrium outcome of salience bias may be that investors 'reach for yield' (Bordalo, Gennaioli, and Shleifer, 2016). If financial service providers also shroud risks (Gabaix and Laibson, 2006), then investors may suffer welfare losses (Inderst and Ottaviani, 2009, 2022).

We capitalize on blockchain records to provide supporting evidence of salience bias in investor preferences. Using the investor-level transactions data across a cross-section of yield farms that compete for investor flows based on salient farm yields, we show that investors are attracted to farms with high salient yields, although they turn out to be riskier ex-post. Thus, we document reaching for yield in decentralized financial markets, even in the absence of financial intermediaries and related agency conflicts. Reaching for yield has been documented in the corporate bond (Becker and Ivashina, 2015; Chen and Choi, 2023), mutual fund (Choi and Kronlund, 2018), money market fund (Kacperczyk and Di Maggio, 2017; Gomes, Peng, Smirnova, and Zhu, 2022), asset-backed securities (Efing, 2020), housing (Korevaar, 2023), and structured product markets (Célérier and Vallée, 2017; Vokata, 2023).

Yield farming is a complex and opaque investment strategy. Thus, we closely relate to the literature on complex structured finance. For example, Henderson and Pearson (2011) suggest that 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. For work on complex securities and structured products, see also Carlin (2009); Carlin and Manso (2011); Carlin, Kogan, and Lowery (2013); Griffin, Lowery, and Saretto (2014); Sato (2014); Amromin, Huang, Sialm, and Zhong (2018); Célérier, Liao, and Vallée (2022); Calvet, Célérier, Sodini, and Vallée (2023); Vokata (2021, 2023); Gao, Hu, Kelly, Peng, and Zhu (2023).

In a significant departure from prior work, we study complex financial products offered through smart contracts operating on a blockchain without centralized financial intermediaries who may drive security design to influence sales. The advantage of our study is that we observe the chain of all transactions at the farm and farmer level. This is in stark contrast to the existing literature on complex securities, which bases its evidence on prices or transactions in primary markets. That feature of our data also enables us to understand investor mistakes (Campbell, 2006; Agarwal, Ben-David, and Vincent, 2017), how investors learn, and how information disclosure and lower perceived complexity change their behavior.

More broadly, our work is related to the emerging literature on decentralized finance.(e.g., Cong, Tang, Wang, and Zhao, 2022; Cong, Harvey, Rabetti, and Wu, 2022; Cong, He, and Tang, 2022) To our knowledge, this is the first empirical study of the risk and return characteristics of yield farming strategies using a hand-collected data set from PancakeSwap. Several studies investigate the properties of automated market makers (AMM) with the constant product model adopted by major decentralized exchanges (DEXs, e.g. Angeris, Kao, Chiang, Noyes, and Chitra, 2021; Aoyagi, 2021; Capponi and Jia, 2021; Han, Huang, and Zhong, 2021; Foley, O'Neill, and Putnins, 2022; Hasbrouck, Saleh, and Rivera, 2022), or focus on strategic trading and liquidity provision (Lehar and Parlour, 2024; Park, 2023; Fang, 2023; Li, Naik, Papanicolaou, and Schönleber, 2024). Appendix Table A.1 illustrates how we differ from these studies. Our main distinction is to exploit wallet-level data and quasi-natural experiments in the yield farm ecosystem to understand channels of yield chasing behavior.

3 Conceptual framework

Yield farming allows investors to passively earn income for their liquidity provision to DeFi platforms. Intuitively, it is similar to securities lending, with the distinctive feature that smart contracts, which operate on permissionless blockchains, automatically execute transactions without involvement of intermediaries. See Appendix A for institutional details.

In practice, yield farming is complex, both in execution and in payoffs. Figure 3 provides a heuristic illustration of the yield farming mechanism in PancakeSwap, the second largest spot decentralized exchange (DEX) offering cryptocurrency exchange services. Figure 3 illustrates that yield farming involves two sequential and independent investment decisions.

First, an investor can passively earn income by providing liquidity to one or several among a large cross-section of liquidity pools. Each pool is defined by a pair of cryptocurrency tokens (USDT-ETH in our example). As liquidity providers, investors 'stake' (i.e., deposit) the pair of cryptocurrency tokens in equal dollar amounts to a liquidity pool. The liquidity provision is certified through the award of a liquidity token, also known as LP token.

Investors are compensated for their liquidity provision through trading fees, which are collected from third party traders who buy and sell USDT–ETH. The trading fees are paid in the pool's currency tokens, i.e., USDT vs. ETH, and amount to 0.25% of a pool's trading volume. Of that amount, 0.17% is paid out to liquidity providers, and 0.08% is paid as a reward to the PancakeSwap main staking contract.

Second, investors can passively earn yield by staking the LP token to a yield farm that is exclusively linked to one liquidity pool (e.g., USDT–ETH). Farm yields are paid in a currency called CAKE, PancakeSwap's native governance token. In this Decentralized Autonomous Organization, CAKE token holders can influence the governance of the PancakeSwap ecosystem by casting votes on the future development of the platform or the reallocation of yields across farms. CAKE ownership also provides rights to participate in services such as non-fungible token (NFT) giveaways or other PancakeSwap lotteries.

CAKE tokens are continuously issued by PancakeSwap's main staking contract with creation of each BSC block. The amount of CAKE tokens allocated to yield farms may vary across farms and over time, as determined by the votes of the aggregate CAKE ownership. PancakeSwap also uses a fraction of the revenue it receives from third party trading fees to continuously buy back and burn (i.e., destroy) CAKE to minimize the currency's dilution.

Based on the complicated chain of transactions described in Figure 3, the total gross return to yield farming between day t and $t+h$, $R_{t,t+h}$, comes from two sources associated with liquidity mining, $R^{\ell}_{t,t+h}$, and the staking of LP tokens to a yield farm, $R^{f}_{t,t+h}$, such that:

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

3.1 Liquidity provision

A liquidity provider must stake a pair (A, B) of cryptocurrency tokens (e.g., USDT and ETH) in equal dollar amounts. This implies that the number (a_t, b_t) of tokens to be pledged is determined by market prices (P^A_t, P^B_t) through the relation $a_t \cdot P^A_t = b_t \cdot P^B_t$.

A pools' aggregate liquidity L_t is characterized by the value of the aggregate number of staked tokens $\alpha^A_t = \sum a_t$ and $\alpha^B_t = \sum b_t$, such that:

$$L_t = \alpha^A_t \cdot P^A_t + \alpha^B_t \cdot P^B_t. \tag{2}$$

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:

$$\begin{aligned} R^{\ell}_{t,t+h} &= L_{t+h}/L_t + Trading\ Fee\ Return_{t,t+h} \\ &= \frac{\alpha^A_{t+h} \cdot P^A_{t+h} + \alpha^B_{t+h} \cdot P^B_{t+h}}{\alpha^A_t \cdot P^A_t + \alpha^B_t \cdot P^B_t} + Trading\ Fee\ Return_{t,t+h}. \end{aligned} \tag{3}$$

Growth in the value of liquidity, L_{t+h}/L_t, is driven by fluctuations in market prices (P_t^A, P_t^B) and by fluctuations in the pool's token composition (α_t^A, α_t^B). Both are pinned down by the constant-product automated market maker (AMM) technology hardwired into liquidity pools. See, e.g., Lehar and Parlour (2024), Capponi and Jia (2021), Park (2023), for details.

The composition of a pool's liquidity changes because third party traders buy or sell tokens A and B, say USDT and ETH. The constant-product AMM technology defines the terms of trade by imposing that, at each point in time, the tokens' product must equal a constant k, i.e., $k = \alpha_t^A \alpha_t^B = \alpha_{t+h}^A \alpha_{t+h}^B$. In other words, the terms of trade are defined through an isoquant curve whose value is determined by aggregate liquidity provision.[1] The constant-product AMM technology also drives price fluctuations, since it imposes, for all t, that the products of price and quantity have to equalize across assets, i.e., $\alpha_t^A P_t^A = \alpha_t^B P_t^B$.

Thus, liquidity providers are exposed to risks associated with joint changes in token prices and token composition. First, in exchange for their liquidity provision, investors receive LP tokens to certify their partial ownership in the pool. While the fractional ownership stays constant, the number of each token that can be claimed at redemption may change with the change in pool composition due to third-party trading. Second, the change in token composition leads to mechanical price changes driven by the constant-product AMM.

In Appendix B, we explicitly show that the growth in liquidity value can be rewritten as a sum of two distinct components that are uniquely functions of prices:

$$L_{t+h}/L_t \quad = \quad \underbrace{\left(\frac{1}{2}R_{t,t+h}^A + \frac{1}{2}R_{t,t+h}^B\right)}_{\text{capital gain}} - \underbrace{\frac{1}{2}\left(\sqrt{R_{t,t+h}^A} - \sqrt{R_{t,t+h}^B}\right)^2}_{\text{impermanent loss}}, \tag{4}$$

where $R_{t,t+h}^A = P_{t+h}^A/P_t^A$ and $R_{t,t+h}^B = P_{t+h}^B/P_t^B$ denote the gross returns of tokens A and B, corresponding to USDT and ETH in our example.

The first term, which we call "capital gain," is the equally-weighted gross return from tokens A and B. The second term describes investors' risk exposure, referred to as impermanent loss. Intuitively, the impermanent loss corresponds to the difference between the return from liquidity provision L_{t+h}/L_t 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 divergence in token returns. Importantly, they are strictly negative and expose investors to significant downside risk analogous to a short volatility exposure (Aigner and Dhaliwal, 2021). See Figure A.5 and 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. PancakeSwap v1 (v2) charges a trading cost equivalent to 20 (25) basis points (bps) of trading volume. Part of that (17bps) 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

$$\textit{Trading Fee Return}_{t,t+h} = c \cdot \left((I_t/L_t)V_{t,t+h}\right)/I_t = c \cdot V_{t,t+h}/L_t. \tag{5}$$

[1] New liquidity provision or redemption can change k, and, therefore, the curvature of the isoquant curve.

3.2 Yield farming

LP tokens may be staked to the main staking contract (yield farms), which offers yield y_t as additional passive income. Yields are paid in CAKE, PancakeSwap's governance token.

The annualized yield is implicitly defined through a non-linear function that depends on (a) the number of CAKE tokens created through new block validation 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 as reward for farming; (d) the total liquidity staked to the farm, L_t^{staked}; and (e) the price of CAKE, P_t^{CAKE}.

Forty CAKE tokens are created with each blockchain validation, which lasts about 3 seconds. This implies that around 28,800 blocks are created per day. Given 365 days in a year, the annualized yield from farming is, therefore, given by:

$$y_t = C \times \left(\frac{m_t}{M_t}\right) \times \left(\frac{P_t^{CAKE}}{L_t^{staked}}\right), \tag{6}$$

where $C = 365 \times 28,800 \times 40$. Since CAKE tokens may be allocated to activities other than yield farming, the aggregate multiplier does not have to equal the sum of all yield farm multipliers, i.e., $M \neq \sum_k m^k$, where k corresponds to the number of farms. The realized farm yield between t and $t+h$, from the perspective of a USD investor, is thus defined as:

$$R^f_{t,t+h} = P^{CAKE}_{t+h} \sum_{n=1}^{h} \left(\frac{y_{t+n-1}}{P^{CAKE}_{t+n-1}}\right)\left(\frac{1}{365}\right). \tag{7}$$

3.3 Aggregation: Frictionless benchmark

The aggregate h-period return to yield farming is thus composed of four components associated with capital gains, impermanent losses, trading fees, and realized farm yields:

$$\begin{aligned} R_{t,t+h} \quad = \quad & \underbrace{\left(\frac{1}{2}R^A_{t,t+h} + \frac{1}{2}R^B_{t,t+h}\right)}_{\text{capital gain}} - \underbrace{\frac{1}{2}\left(\sqrt{R^A_{t,t+h}} - \sqrt{R^B_{t,t+h}}\right)^2}_{\text{impermanent loss}} + \underbrace{c \cdot V_{t,t+h}/L_t}_{\text{trading fee revenue}} \\ + \quad & \underbrace{P^{CAKE}_{t+h} \sum_{n=1}^{h} \left(\frac{y_{t+n-1}}{P^{CAKE}_{t+n-1}}\right)\left(\frac{1}{365}\right)}_{\text{realized farm yield}}. \end{aligned} \tag{8}$$

3.4 Impact of trading frictions

Table A.2, which breaks down the chain of transactions for a hypothetical yield farming strategy, shows that harvesting yields from PancakeSwap involves a chain of 12 transactions (excluding step 1 and 14 in Table A.2). A full round-trip transaction involves three types of costs associated with gas fees, trading fees, and price impact (see Appendix 5.5 for details). 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 pay gas fees for 10 of them. The average round-trip gas fee is estimated to be \$3.45 in our sample period.

Since gas fees are flat overhead costs, they are more costly for small stake investments and frequent rebalancing. Thus, they are especially detrimental to smaller retail investors, who have a tendency to rebalance too frequently (Odean, 1999). Since gas fees grow linearly with each additional yield farm, there is also less benefit from diversification across farms. As a result, gas fees encourage larger and more concentrated investments, which may not be optimal for financially unsophisticated investors.

For example, a \$1,000 investment would lose ≈35 bps in gas fees for a round-trip transaction, and 35 bps per week for weekly rebalancing. A diversification strategy across 10 farms would incur a per period cost of $10 \times 3.45 = \$34.5$, which, for a \$1,000 investment, is more than the typical hedge fund performance fee, not considering hurdle fees or water marks.

Besides gas fees, investors are charged a trading fee of 0.25% (proportional to trading volume) per transaction. Since one round-trip transaction includes the buying and selling of tokens at intermediate steps, yield farmers lose at least an additional 0.50% of their initial investment. The selling of CAKE tokens at redemption also requires a proportional trading fee of 0.25%. See Appendix 5.5 for more details.

The third transaction cost arises through price impact. We characterize a price impact function $\lambda(f)$, where f denotes the ratio of the investment amount I_t to the value of the liquidity L_t, i.e., $f = I_t/L_t$. Panels (a) to (c) of Figure A.6 illustrate how price impact is increasing in the size of an investment relative to a pool's liquidity. Considering both trading fees and price impact, the growth in liquidity value reduces to:

$$L_{t+h}/L_t = (1 - 0.0050)\lambda(f)\left[\left(\frac{1}{2}R^A_{t,t+h} + \frac{1}{2}R^B_{t,t+h}\right) - \frac{1}{2}\left(\sqrt{R^A_{t,t+h}} - \sqrt{R^B_{t,t+h}}\right)^2\right]. \tag{9}$$

We emphasize another indirect channel that negatively affects yield farming performance. Equation (6) highlights a negative relation between a yield farm's aggregate liquidity and the offered farm yield. We provide empirical support for that relation in Figure A.1. 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.

3.5 Investor mistakes and aggregation with frictions

Farm yields are strictly non-negative and yield farms are built on the same blockchain as liquidity pools. Thus, in the absence of lock-up periods, the staking of LP tokens is always a dominating strategy and the optimal staking ratio k^* should equal one. Because all transactions are observed on the blockchain, we can identify when investors do not stake their LP tokens into yield farms. We consider staking ratios below one to be a mistake.

Including all trading frictions, we quantify the returns to yield farming as follows:

$$R_{t,t+h}^{friction} = (1-0.0050)\lambda(f)\left[\underbrace{\left(\frac{1}{2}R_{t,t+h}^{A} + \frac{1}{2}R_{t,t+h}^{B}\right)}_{\text{capital gain}} - \underbrace{\frac{1}{2}\left(\sqrt{R_{t,t+h}^{A}} - \sqrt{R_{t,t+h}^{B}}\right)^{2}}_{\text{impermanent loss}}\right]$$
$$+ \underbrace{c \cdot V_{t,t+h}/L_{t}}_{\text{trading fee revenue}} + (1-0.0025)\,k^{*}\underbrace{\left[P_{t+h}^{CAKE}\sum_{n=1}^{h}\left(\frac{y_{t+n-1}}{P_{t+n-1}^{CAKE}}\right)\left(\frac{1}{365}\right)\right]}_{\text{realized farm yield}} - \frac{Gas_{t,t+h}}{I_{t}}. \quad (10)$$

3.6 Yield farm flows

In our analysis, we examine two measures of yield farm flows. First, we measure farm flows using LP token growth, i.e., $Flow_{t,t+h} = (\#LP\ tokens_{t+h}/\#LP\ tokens_{t}) - 1$. Because this measure is insensitive to price fluctuations, it is conceptually similar to the use of net fund flows adjusted for price growth used in the mutual fund literature (e.g., Sirri and Tufano, 1998; Coval and Stafford, 2007). Second, we measure farm flows using the dollar growth in pool liquidity, i.e., $Flow_{t,t+h} = (L_{t+h}/L_{t}) - 1$.

4 Building yield farm and yield farmer data

We assemble a novel data set on liquidity pools and yield farms listed on PancakeSwap by tracing information on the Binance Smart Chain. Our data include the full history of prices, transactions, token shares, liquidity provision, and yield farm multipliers.

4.1 Farms and yields

We consider all contract addresses of liquidity pools with a corresponding yield farm stored in PancakeSwap's main staking contract from their inception on September 23, 2020 to July 31, 2022. With these addresses, we reconstruct, from the blockchain, the daily time series of farm yield multipliers. We consider only active farms with a non-zero yield multiplier.

Farm yields are a function of aggregate farm liquidity. We, therefore, source each pool's token balances α_t^i and prices P_t^i to compute aggregate pool liquidity given by $L_t = P_t^A\alpha_t^A + P_t^B\alpha_t^B$ (See Equation 2). Next, we collect each pool's supply of LP tokens and their staking ratios to compute aggregate farm liquidity defined as $L_t^{staked} = L_t \cdot (\#\text{ staked LP tokens/Total }\#\text{ of LP tokens})$.

We impute offered farm yields using Equation (6). We verify their accuracy by collecting offered farm yields from PancakeSwap's homepage[2] at midnight Greenwich Meridian Time (GMT) on October 11, 2021. We manually verify that the multipliers collected from the main staking contract are identical to those advertised on PancakeSwap's web interface.

[2] See https://PancakeSwap.finance/farms.

Figure A.2 reports the relation between our imputed farm yields and those publicly listed by PancakeSwap. Nearly all observations are closely aligned with the 45-degree line. A linear projection of the imputed on the listed farm yields obtains a slope coefficient of 1.002 with an R^2 of 1.00. This strongly supports the validity of our data building procedure.

4.2 Prices, trades, and transaction costs

In a liquidity pool (e.g., ETH–BNB), one token of the cryptocurrency pair is considered a token of interest (e.g., ETH). Its price is typically expressed in terms of a numeraire token (e.g., BNB). We source daily end-of-day GMT prices, P_t^i, of the tokens of interest.

To find the prices of the numeraire token (BNB), we first use the native historical quote function on PancakeSwap. This pins down the historical exchange rate between BNB and Binance-Peg Tether (USDT), a stablecoin pegged to the US dollar. We then convert USDT to U.S. dollars using the USDT price from CoinMarketCap. This allows us to compute the U.S. dollar h-period trading volume $V_{t,t+h}$ as the daily sum of all trades.

We source gas fees from Bitquery, a proprietary data vendor specialized in blockchain data services covering BSC and other blockchains. Gas fees differ across functions executed by smart contracts. To accurately account for transaction costs in computing the performance of yield farming strategies, we first identify the transactions that incur gas fees (see Table A.2) and compute their average daily gas fee in U.S. dollars. Next, we compute the round-trip cost of gas fees by summing the average fee across all relevant transactions.

4.3 Yield farmers

We collect transaction data for all LP tokens from the transaction logs of BscScan[3], a freely-accessible analytics platform on BSC, and reconstruct each wallet's historical token holdings. Transactions that involve a user's deposit of cryptocurrency to a liquidity pool in exchange for LP tokens are represented as LP token transfer from the null address (0x000...000) to the user's wallet address. Transactions in which a user stakes/unstakes their LP tokens in a yield farm are captured as a token transfer to/from the active main staking contract. Redemptions of LP tokens at a liquidity pool in exchange for the underlying tokens are represented as a LP token transfers to the address of the LP token.

We restrict our analysis to active accounts. We eliminate wallet addresses that are not associated with PancakeSwap smart contracts and accounts with more than 100,000 trades, since those wallets may camouflage yield aggregators or automated passive strategies. Relatedly, we remove positions lower than \$1 at the beginning of the holding period because they are below the average exit cost (see Figure 2), possibly distorting the analysis. In addition, we omit wallet addresses that have transacted LP tokens with third party smart contracts outside PancakeSwap, since the study of staking across multi-platform investment strategies is beyond the scope of our study. For accounts with positive end-of-sample LP token balance, we assume that farmers liquidate their open positions on July 31, 2022.

[3] See `https://bscscan.com/`.

We merge each transaction with information on token prices and offered farm yield using the nearest end-of-day prices by block height difference. For each wallet, we also compute the number of invested farms (*No. Farms*) and liquidity pools (*No. Pools*). We define *Efficiency* at the wallet level as the duration of staking relative to the duration of liquidity provision (*Time Staked*/*Time in Liquidity Pool*), averaged across liquidity pools. Third, we define *Staked Balance* and *LP Balance* as the time-weighted average balance for staking and liquidity provision. For these calculations, we use the nearest end-of-day price from the beginning of each holding period and weight balances by the length of the holding period.

We define *Offered Farm Yield* at the yield farmer level as the time-weighted average offered yield at the beginning of each holding period. Finally, we calculate a farmer's *Average Daily Return* as the time-weighted geometric average of their holding period returns. We compute all return components as described in Section 3, making the simplifying assumption that offered yields are harvested daily without reinvestment.

Yield farmers may split their investments across multiple wallets. Hence, measures such as *No. Farms*, *Staked Balance*, and *LP Balance* could be underestimated. However, yield farmers are unlikely to systematically manage multiple wallets since there are no monetary benefits and transaction costs significantly increase. This bias, which is not central to our analysis, could be examined in future research using wallet clustering algorithms.

4.4 The final sample

Our final sample contains 262 unique yield farms that were active between the inception of PancakeSwap on September 23, 2020 and July 31, 2022. At the investor level, we analyze 439,639 (446,227) unique wallets which hold 6,183,222 (5,975,858) positions for the return (flow) analysis. Appendix C provides a detailed account of our data cleaning and construction procedure.

Panel (a) of Figure 4 illustrates the number of active farms (right axis). The cross-section varies over time since new farms may be listed or delisted. The total number of active farms increases quickly from inception of PancakeSwap to a peak of 160 farms in July 2021.

The left axis in Panel (a) of Figure 4 plots Total Value Locked (TVL) in active farms, i.e., the aggregate amount of LP tokens staked to yield farms. Yield farming at PancakeSwap has experienced extraordinary growth, with TVL surpassing $7 billion in May 2021. Analogously to the boom-bust cycles experienced by Bitcoin and other cryptocurrencies, TVL dropped sharply following its peak and experienced renewed momentum.

Importantly, TVL remained subdued until early 2021. As we show in Panel (b) of Figure 4, the consequential increase in liquidity provision coincides with PancakeSwap becoming more prominently researched in Google (left axis). Simultaneously, the number of active farmers jumps sharply (right axis). For that reason, we restrict our main analysis to start on March 1, 2021 to increase the stability of our estimations and avoid noisy inferences.

5 Evidence

We first provide new stylized facts on yield farms and farmers. We then describe the trading behavior of yield farmers and examine the risk and return characteristics of yield farming.

5.1 Stylized facts about yield farms

Yield farms exhibit three important properties for our analysis of reaching for yield. First, they promise extraordinary high yields with cross-sectional heterogeneity in earnings potential. Second, promised yields are salient and conspicuously displayed as headline rates, while information on risks and historical performance is hidden and difficult to find. Third, yield farms appear as simple and engaging platforms but involve a high degree of complexity.

We report in Table 1 a snapshot of yield farms on July 31, 2022. Each farm features a unique pair of cryptocurrency tokens. Panel A shows the ten largest farms in terms of TVL. The largest farm draws from $178.28 million TVL in the USDT–BUSD pool. In Panel B, we show that the leading farm in terms of earnings potential offers an annualized yield of 357.92% for TVL of $1.72 million in the BTCST–WBNB liquidity pool. Yield farms feature considerable cross-sectional heterogeneity in terms of liquidity and earnings potential. For example, the rankings in Table 1 show that TVL ranges from $0.12 million to $178.28 million (Panel A), while yields range from 0.30% to 357.92% (Panel B).

In Figure 5, we plot the time-variation of the median farm yield together with its cross-sectional distribution. To be precise, we plot the total offered yield which is saliently disclosed to investors on PancakeSwap's webpage and referred to as the annualized percentage return (APR). This includes both the offered yield (Equation (6)) and the trading fee yield estimated using the previous day's trading volume. The median farm yield is often higher than 50% and the average is 76.34%. In addition, there is significant variation in dispersion of farm yields, as underscored by the fluctuations of the interquartile range of the yield farm distribution. Such rich variation in yields across farms and time together with transparency on blockchain transactions provides an opportunity for better understanding the motivations behind liquidity provision to yield farms and the performance of yield farming.

Yields are salient to investors and marketed as headline rates that look attractive, especially in a low interest rate environment. In Appendix Figure A.3, we provide an example of PancakeSwap's user interface. The main information in the foreground relates to the total offered yield (i.e., the APR), the yield multiplier and the pool's liquidity.

In contrast, it is difficult to find information about the computation of annualized returns or the meaning of yield multipliers. Moreover, it is difficult to find information about the return decomposition. There are hidden downside risks associated with impermanent losses, and hidden costs due to the price impact of large trades, also known as slippage.

The user interface of PancakeSwap is engaging because it displays cartoons, rockets and emojis. This gamification of an investment platform makes yield farming look like a simple application. It is, however, a complex investment strategy, both in terms of payoffs and execution. The payoffs to yield farming depend on three underlyings: the two cryptocurrencies in the liquidity pool and PancakeSwap's governance token CAKE, which is paid as a reward for yield farming. Furthermore, the payoffs feature significant non-linearities, epitomized by the impermanent loss function. Finally, a round trip in yield farming is complex to execute, since it involves a chain of up to 14 transactions (see Appendix Table A.2).

5.2 Determinants of farm yields

In Equation (6), we characterize the offered farm yield as a function of five components. Among these, one is mechanically related to the continuous CAKE token issuance (c), one depends on the aggregate CAKE price (P_t^{CAKE}), and one depends on farm-specific liquidity ($L_{i,t}^{staked}$). These factors are outside the influence of CAKE owners. The farm-specific multiplier $m_{i,t}$ defines the allocation of CAKE tokens to a farm. The multiplier M_t defines the aggregate distribution of CAKE tokens. We validate that all components are strongly correlated with the level of offered farm yields and that they have the correct sign.

CAKE owners can vote on changing the farm-specific multiplier $m_{i,t}$ to increase or decrease the CAKE token allocation. Since increasing the farm multiplier increases the offered farm yield, CAKE owners can influence the reallocation of rewards across farms and attract liquidity to a liquidity pool of choice. Thus, the ability to change the yield multiplier $m_{i,t}$ equips CAKE owners with centralized decision power on the amount of passive earnings potential, which, in our opinion, goes against the spirit of decentralized financial services.

We want to examine the determinants of yield changes that are associated with decisions to change the yield multiplier, controlling for all common variation (e.g., M, L, P^{CAKE}). In Table 2, we, therefore, isolate the impact of yield changes that come from the active decision of farm governors (i.e., the owners of CAKE tokens). We examine the relation between the change in yield that is driven by platform governance ($\Delta y_{i,t+1}^m = y_{i,t} \times \frac{\Delta m_{i,t+1}}{m_{i,t}}$) and various components of the yield farming return performance over the previous seven days, i.e., capital gains, impermanent loss, trading fees, realized yields, and farm liquidity.

Columns (1) and (2) of Table 2 suggest that yields are increased when past trading fees are high, and decreased when past realized yields are high. This result holds both with and without day fixed effects that absorb common movements across farm yields due to, for example, the price of CAKE.

In columns (3) and (4) of Table 2, we find that farms are more likely to be delisted when their liquidity or trading fee revenues are low. Overall, this evidence is consistent with the idea that the offered farm yield is an instrument to make the strong farms stronger and the weak farms weaker. Thus, offering yields is a mechanism to enhance the long-term viability of the yield farm platform by channeling liquidity to a subset of farms.

5.3 Evidence on lack of investor sophistication

Several infrastructure developments of PancakeSwap enable us to examine trading behavior. First, PancakeSwap upgraded the technological and security features of its smart contract design on April

24, 2021, migrating from 'PancakeSwap v1' to a new version 'PancakeSwap v2'. 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 encouraged to switch their liquidity provision from v1 to v2, but had to implement the switch themselves. The switch is strictly dominating, because migrating liquidity to the new version delivers higher staking rewards than in v1, alongside lower transaction costs.

In Panel (a) of Figure 6, we show the amount of outstanding assets in the old version of the platform. This figure shows that the migration of funds is sluggish, which could be a sign of investor inattention or inertia. Importantly, even after 100 days, a significant amount of liquidity remains in the liquidity pools associated with the old version.

A second update occurred on April 20, 2022, when PancakeSwap migrated its staking functionality to a new contract. Users were encouraged in advance, through Twitter and other PancakeSwap platforms, to migrate their assets.[4] Migrating is again preferred because assets in the old staking contract would stop earning yields. Panel (b) of Figure 6 shows a similar pattern in that many users remain staked in the obsolete staking contract even 100 days after the migration, missing out on potential yield income in that period. This phenomenon is similarly a sign of investor inertia, inattention, or of their lack of sophistication.

More evidence on investor behavior comes from staking ratios, defined as the ratio of LP tokens staked in yield farms to the aggregate amount of LP tokens minted to certify liquidity provision. Remember that yield farmers sequentially provide liquidity to pools and then to farms. Implementing both transactions is strictly dominating liquidity provision alone, since earning CAKE through farming is always superior to leaving money on the table. Thus, we would expect the staking ratio to be equal to one at all times.

Figure 7 shows that the median staking ratio is below one most of the time. The 10th (25th) percentile of the distribution drops as low as 20.56% (40.51%). This further suggests that some investors are financially unsophisticated. We caveat this interpretation because of the possibility that investors stake their LP tokens in third-party yield farm aggregators.

Table A.4 in the Appendix shows that staking ratios increase with experience. We regress the staking ratios on indicator variables that are one for the 3rd (4th, 5th, >5) farm investment and zero otherwise. The constant, which captures the baseline for the first two farms, indicates that the average staking ratio is 62.39%, based on the linear probability model in column (1). The staking ratio significantly increases with every subsequent farm investment, suggesting that investors learn over time.

In Panel A of Table 3, we provide farmer-level statistics. The average yield farmer invests in 2.64 farms, has a holding period of 30.92 days, and has $7,732 invested. However, the average staking ratio is only 0.8422. This suggests that many farmers miss out on farming opportunities, possibly due to the complex nature of the trading strategy.

In Panel B of Table 3, we separate the farmers into quintiles based on their average *Investment Size*. There is significant cross-sectional dispersion in size among PancakeSwap users. For instance, the average investment in the lowest quintile is only $10.96, whereas that of the highest quintile is $37,738. Thus, many yield farmers have small investment stakes.

[4] See, for example, https://twitter.com/PancakeSwap/status/1385463720835379201.

We observe that *Investment Size* is positively correlated with the staking ratio, ranging from 0.6211 to 0.9629 between the lowest and highest quintiles. This suggests that smaller yield farmers are more likely to leave money on the table. But even in the highest quintile do we see significant evidence of investor mistakes, given an average staking ratio that is far from one. Since the average farm yield ranges between 94.28% and 121.69% across quintiles, investors face non-trivial opportunity costs. Investors also have short holding times, with a holding period ranging between 9.8 and 61.31 days across quintiles.

Investment performance is non-linear across quintiles with the highest daily return of 20bps in quintile 4. This echoes our discussion in Section 3.4 that both large and small investments could generate sub-optimal performance, due to transaction costs and price impact.

Evidence from DappRadar[5] indicates that PancakeSwap registered 435,130 active users on October 24, 2021, in contrast to 47,730 active users recorded for Uniswap. The trading volume in PancakeSwap was about $1.2 billion on that day, which implies that the average yield farmer in PancakeSwap traded $2,757. This suggests that many investors in PancakeSwap are small retail investors, consistent with our evidence in Table 3.

Survey evidence further supports the view that yield farmers may not be financially sophisticated. CoinGecko, a data provider, questioned 1,347 cryptocurrency investors about yield farming in August 2020 (CoinGecko, 2020). According to the survey, 79% of yield farmers claim to understand the risks and rewards of yield farming to a reasonable extent. However, about 40% of them report that they could not read smart contracts to verify potential yield vulnerabilities or scams. In addition, 33% of yield farmers are unfamiliar with the meaning of impermanent loss, implying that they take risks which they don't understand.

5.4 Yield farming performance without frictions

In Table 4, we assess the value-weighted performance of yield farming strategies using aggregate pool liquidity as the weighting factor. We compute returns in excess of the 3-month U.S. Treasury bill rate from the perspective of a U.S. investor and ignore transaction costs. Panel A (B) reports results at the daily (weekly) trading frequency with 518 (74) observations.

We find that, prior to transaction costs, yield farming is profitable during our sample period. The value-weighted index strategy delivers a daily return of 0.15%. This is about twice as large as the returns to a strategy that focuses only on liquidity mining (0.07%) or on a buy-and-hold strategy in the same pairs of cryptocurrency tokens associated with the liquidity pools (0.07%). All three strategies deliver negatively skewed performances, with a non-trivial amount of excess kurtosis, negative serial correlation, and exhibit a daily volatility of about 3.6%. Results for a weekly trading frequency are qualitatively similar.

In Figure 8, we report the performance for each the four components (capital gains, impermanent losses, trading fees, farm yields) after sorting yield farms into quintiles based on the magnitude of their average in-sample offered yield. Panel (a) shows that the average realized yield, which strongly correlates with the offered yield, increases monotonically across quintiles from about 2bps

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

in Q1 to about 40bps in Q5. Panel (b) of Figure 8 shows that trading fee revenue is smaller than other components and more similar across quintiles.

In Panel (c), we illustrate capital gains. While capital losses are largest for farms offering high yields, these are insignificantly estimated. In contrast, impermanent losses, which are significantly estimated, as shown by the 95% confidence bounds, are always negative and monotonically decreasing with the headline yields, as shown in Panel (d). Taken together, this evidence at the farm level suggests that high yield farms' tokens generate the lowest returns and the largest impermanent losses.

The evidence that farms with the highest headline yields perform worst ex-post raises concerns about retail investor protection for three reasons. First, yield farms compete for liquidity by offering high yields. Second, high yields are salient to investors who appear to be unsophisticated. Second, impermanent losses are shrouded, yet they significantly contribute to yield farming underperformance. Bordalo, Gennaioli, and Shleifer (2016) show that, in such an environment, reaching for yield behavior may be an equilibrium outcome.

To better assess the risk-return trade-offs, we standardize the return performance by the standard deviations and report in Table 4 Sharpe ratios for all investment strategies. These measures suggest risk-return trade-offs of yield farming that are comparable but higher to that of the S&P 500 (which has a daily Sharpe ratio of 0.03 in our sample period), with values ranging from 0.0209 for buy-and-hold strategies to 0.0405 for yield farming.[6] Thus, without accounting for frictions, yield farming appears to be profitable and to deliver superior performance to the S&P 500, according to Sharpe ratios.

We also report alphas estimated using the three-factor cryptocurrency return model of Liu, Tsyvinski, and Wu (2022), in addition to BNB, the native token of the BSC smart chain. Their framework suggests that a three-factor model with cryptocurrency market, size, and momentum factors prices the cross-section of cryptocurrency returns. Thus, we assess the risk-adjusted yield farming performance relative to this three-factor+BNB cryptocurrency benchmark. We find that the daily yield farming alpha is, on average, 0.02%. Because of the short and volatile sample period, this alpha is estimated with a t-statistic of only 0.6822. The alphas of buy-and-hold investments and liquidity mining are negative, emphasizing that the positive yield farming performance is driven by farm yield and trading fee revenue.

5.5 Yield farming performance with frictions and investor mistakes

We next consider the impact of investor mistakes by comparing the performance of yield farming to that of liquidity mining. Panel (a) of Figure 9 shows that investors who do not fully stake their LP tokens into yield farms perform worse within each quintile. This effect is especially pronounced for the farms with the highest headline rates. Table 3 documents that smaller investors are more likely to make mistakes (i.e., staking ratios below one). Thus, they are more likely to leave money on the table and underperform. Detailed statistics are reported in Panel A of Appendix Table A.7.

We further assess the impact of trading frictions on yield farming performance, including gas fees, trading fees, and price impact. For that purpose, we assume a holding period of 10 days, or that

[6] There are fewer observations for the S&P500 because DeFi markets are continuously open for trading.

1/10th of the investors rebalance their portfolio each day. This lies within the mean and median holding periods across yield farmers (see Table 3). We choose an initial investment of \$1000, which is bounded by the mean and median investment amount in our sample. Finally, we approximate the staking ratio of the average investor using the average daily observed farm-level staking ratio.

Panel (b) of Figure 9 compares the yield farming performance with trading frictions and investor mistakes to that of the frictionless benchmark. Transaction costs unilaterally lower the risk-adjusted return performance across all yield quintiles. For example, the risk-adjusted return decreases by 7bps from 0.07% (−0.01%) to 0.00% (−0.08%) for Q1 (Q5). That downward adjustment is further amplified by investor mistakes such that, for Q5, the daily alpha decreases from −0.01% to −0.21% (see Table A.7, Panel A). Figure A.8 and Panel B in Table A.7 provide qualitatively similar results at the weekly trading frequency, but the downward adjustments are larger in magnitude.

Figure A.9 illustrates robustness of our conclusions by showing similar results under alternative parameter assumptions for the trading frictions. In Panel (a), we first vary the rebalancing duration from 7 to 14 days. We report annualized alphas for a fair comparison across scenarios. Risk-adjusted returns decrease monotonically within each quintile. This is expected, since the multiplicity of transactions needed for a round-trip trade can accumulate to non-trivial amounts for gas and trading fees, especially with frequent rebalancing.

In Panel (b) of Figure A.9, we vary the investment size from \$500 to \$5000. Small size investments are impacted by gas fees, since these are based on flat dollar amounts. This incentivizes larger investment amounts to reduce the dollar cost per investment. However, larger amounts may not be an option for unsophisticated retail investors. Indeed, a large proportion of investors invest less than \$1,000 (see Table 3). On the other hand, large investments relative to the size of the liquidity pool may suffer from price impact due to slippage. In addition, larger investments can endogenously lead to lower farm yields, thereby putting further downward pressure on the investment performance. Hence, we observe hump-shaped performance results within each quintile.

These observations bear implications for diversification and optimal portfolio allocation. A portfolio with 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. We leave such analysis for further research.

In Table 5 we examine the role of trading frictions and investor mistakes at the farmer level. We regress the time-weighted average daily holding period return for each farmer on the average value-weighted displayed farm yield and a set of explanatory variables related to transaction costs and mistakes. In columns (1) to (3), farmer-level returns without frictions are the dependent variable and there is, at best, weak significance by any of the explanatory variables.

In contrast, in columns (4) to (6) of Table 5, which do account for frictions, all variables are strongly significant in explaining daily holding period returns. Investment size is non-linearly related to performance, as underscored by the positive and negative coefficients on investment size and its square. More frequent rebalancing is associated with higher gas and trading fees and lower performance, while higher staking ratios in yield farms lead to better performance since less money is left on the table. The average difference in daily return performance for a staking ratio of zero and one equals 1.93% to 2.04%. These results are robust to the inclusion of investment start and end month fixed effects, which effectively allows for a comparison between investors over similar trading horizons.

Importantly, the coefficient on the average total offered yield is negative and statistically significant at the 5% level. This suggests that farmers who invest in higher-yielding farms underperform by an additional 24bps for every 100% increase in total offered yields. That evidence is consistent with our findings at the farm level (Figure 9) in that the farms with the highest headline rates exhibit the worst risk-adjusted performance. 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, 2023; Bordalo, Gennaioli, and Shleifer, 2016; Vokata, 2023; Gomes, Peng, Smirnova, and Zhu, 2022) and 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).

6 Reaching for yield in decentralized financial markets

The PancakeSwap ecosystem hosts a large cross-section of yield farms that compete for liquidity by offering seemingly attractive investment opportunities while shrouding risks. The detailed account of all wallet transactions registered on the public blockchain provides a unique opportunity to examine whether and how such an environment encourages reaching for yield behavior (Bordalo, Gennaioli, and Shleifer, 2016).

We first examine whether yield farmers adjust their positions in response to changes in farm yields. To that end, we examine the impact of yield changes on liquidity pool flows. We consider both dollar growth and LP token growth, a measure that is similar to net fund flows to mutual funds (e.g., Sirri and Tufano, 1998; Coval and Stafford, 2007).

Equation (6) shows that yields are driven by many factors which are either farm-specific or common to all farms. We would like to isolate the variation associated with the farm multipliers $m_{i,t}$, since changes in multipliers stand out in PancakeSwap (see Figure A.3) and are changed by votes of the platform owners. We also want to avoid capturing fund flows that are driven by multiplier changes to other farms, and therefore restrict our analysis to changes in farm multipliers where the change in the aggregate multiplier M is small.

We identify 511 cases where $\Delta m_{i,t} \neq 0$ and $|\Delta M_t/M_t| \leq 0.15$, among which 50 (461) cases are associated with an increase (decrease) in $m_{i,t}$. We then compare the change of flows into the treated farms with $\Delta m_{i,t} \neq 0$ to those into the non-treated farms with $\Delta m_{i,t} = 0$. Specifically, we plot the difference-in-differences coefficients β_k from a regression:

$$y_{i,t+h} = \alpha + \sum_{k=-7,k\neq-1}^{k=7} \beta_k I_k \times Treatment_i + Event \times FarmFE + DayFE + \varepsilon_{i,t+h},$$

where $y_{i,t+h}$ is defined as either $log(\frac{outstanding\ LP\ tokens_{i,t+h}}{outstanding\ LP\ tokens_{i,t-1}})$ or $log(\frac{\$\ of\ pool_{i,t+h}}{\$\ of\ pool_{i,t-1}})$. Panels (a) and (c) ((b) and (d)) in Figure 10 document significant pool inflows (outflows) on the day that farm multipliers increase (decrease). Token growth (dollar growth) for $\Delta m_{i,t} > 0$ is about 17.94% (19.11%) on day 0, on average, which is economically meaningful.

Platform owners may increase farm multipliers in anticipation of future inflows. To mitigate that concern, we also examine the sensitivity of aggregate fund flows to yield changes that are associated with multiplier changes by peer farms, as reflected in the aggregate multiplier M_t. These shocks need to be large enough to have meaningful impact on M_t and, therefore, $y_{i,t}$. Thus, we identify 4 events where $\Delta m_{i,t} = 0$ with $|\Delta M_t/M_t| > 0.15$. These 4 events are associated with increases in M_t. Since changes in M_t affect all farms simultaneously, we conduct a simple event study without control group. Appendix Figure A.10 confirms the finding that aggregate fund flows are sensitive to changes in yields.

In Table 6, we provide more direct evidence on reaching for yield behavior by testing whether future flows are related to high yield farms. Specifically, we regress farm flows on total offered farm yield, lagged farm performance (*Return*) and the individual components related to capital gains, impermanent losses, trading fees and realized farm fields:

$$\begin{aligned} Flow^j_{t,t+7} = \alpha &+ \beta_1 Total\ Offered\ Yield^j_{t-7,t} + \beta_2 Capital\ Gain^j_{t-7,t} + \beta_3 Impermanent\ Loss^j_{t-7,t} \\ &+ \beta_4 Trading\ Fee^j_{t-7,t} + \beta_5 Realized\ Yield^j_{t-7,t} + \gamma^\top X^j_t + FEs + \varepsilon^j_t, \end{aligned} \tag{11}$$

where j denotes the farm-level index. We include farm and week fixed-effects. The control vector X_t includes lagged flows, log size of the liquidity pools, and farm age.

In column (1) of Table 6, we find a positive and statistically significant relation between $Total\ Offered\ Yield$ and $Flow$. This result is unchanged when we add lagged return performance in column (2). Besides the statistical significance at the 1% level, the coefficient is also economically significant. A farm with a 100% higher $Total\ Offered\ Yield$ is associated with a 6.35% greater increase in fund inflows.

In column (3), we add the four components of lagged return performance and drop $Total\ Offered\ Yield$ due to its high collinearity with $Realized\ Yield$. We find a positive and strongly significant relation between farm flows and lagged trading fees and realized yields. Importantly, these measures are directly observable to investors in the PancakeSwap user interface. This strongly suggests that flows chase past fees and high yields.

The coefficient on *Impermanent Loss* is insignificant, which is consistent with the evidence that information on impermanent losses is challenging to find and difficult to understand, according to survey evidence. Overall, our results suggest that yield farmers chase farms offering higher, more salient yields, but do not seem to internalize past impermanent losses.

6.1 The role of learning in reaching for yield behavior

We next examine reaching for yield at the farmer level. Column (1) in Table 7 reports our baseline result for the relation between flows by investor i to farm j, $Flow^{i,j}_{t,t+7}$, and farm j's offered yield, $Total\ Offered\ Yield^j_{t-7,t}$. The positive and statistically significant coefficient of 0.0255 indicates a positive propensity of reaching for yield. The average farmer provides about 2.55 percentage points more liquidity to a farm if it offers a 100% larger yield.

A significant literature has highlighted the underperformance of yield-seeking strategies (e.g., Henderson and Pearson, 2011; Becker and Ivashina, 2015; Bordalo, Gennaioli, and Shleifer, 2016;

Célérier and Vallée, 2017; Choi and Kronlund, 2018). Similarly, we find that investor funds are more likely to be channeled to higher-yielding farms, which systematically underperform due to greater capital and impermanent losses. The underperformance is especially pronounced for investors who leave money on the table due to their mistake of not staking the LP tokens to earn farm rewards. We are, therefore, interested in understanding whether experience and learning can contribute to mitigating reaching for yield behavior.

In columns (2) to (7) of Table 7, we ask whether proxies for learning and experience can reduce the reaching for yield propensity, defined as the regression coefficient between total offered farm yields and future fund flows. Our three proxies for learning and experience are the amount of the investment (Size), the number of days elapsed since the first yield farm investment (Experience), and the number of farms to which an investor has provided liquidity (# Farms). For all three measures, we create indicator variables equal to one if the variable is above the 75th percentile of the variable's distribution, and zero otherwise.

Columns (2) and (3) provide weak support for investment size playing a role in mitigating reaching for yield behavior. However, columns (4) to (7) show a significant reduction in the reaching for yield propensity based on the interaction terms between the total offered yield and the experience proxies. The most conservative estimations in columns (5) and (7) include farm times week effects, allowing us to control for time-varying farm characteristics and compare high with low-experience farmers within the same farms at different points in time. The magnitude of the coefficients in these estimations suggests that experience can mitigate the reaching for yield propensity by 20% (0.0050/0.0254) to 39% (0.0117/0.0303).

6.2 The role of information disclosure in reaching for yield: Yieldwatch

High yield-seeking behavior has been 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; Vokata, 2023). Much of that research emphasizes the role of complexity and risk shrouding in explaining reaching for yield behavior (e.g., Gabaix and Laibson, 2006; Célérier and Vallée, 2017). A main advantage of the blockchain data is that it allows us to directly test, using natural experiments, whether information disclosure and reduction in complexity can alleviate reaching for yield behavior.

In particular, we rely on the novel setting of Yieldwatch.net, a third-party information platform that selectively discloses information on past performance and return components in exchange for buying Yieldwatch tokens. Launched on March 3, 2021, Yieldwatch Pro, Yieldwatch.net's main service, provides customized information on yield farming. Appendix Figure A.4 provides a screenshot of Yieldwatch Pro's user interface.

Unlike PancakeSwap's main user interface, which provides limited information on farm-level characteristics like yield, size, and multiplier (see Appendix Figure A.3.), Yieldwatch Pro provides a more user-friendly interface with richer information. In addition to information on farm characteristics, YieldWatch Pro breaks down farmers' historical capital gains (also called HODL value), impermanent losses, trading fee revenue, and realized yields for a particular yield farming position. Notably, this information is only available to yield farmers who own Yieldwatch.net's native utility token, called the WATCH token.

We leverage two unique features of YieldWatch Pro to construct shocks to individual investors' information display. First, through the complete transfer history of WATCH tokens available from Binance Smart Chain, we identify WATCH token holders and their balances on each day. Second, an initial farm offering (IFO) of WATCH tokens on March 4, 2021 was designed in such a way that only those investors bidding more than $570 were allocated WATCH tokens. Since that threshold was unpredictable, we can compare yield-chasing behavior among investors just above (treated) and below (control) that cutoff level, using the quasi-random token allocation as a randomized shock to investors' information.[7]

Column (1) of Table 8 provides the baseline result for the reaching for yield propensity based on the regression of individual investors' farm *Flows* on the one-week lagged farm-specific *Total Offered Yield*. To alleviate concerns associated with endogenous token acquisition, we compare the reaching for yield propensity of the WATCH token holders to that of matched control wallets that have a similar farm composition over the previous 180 days (number and types of farms) and TVL deviating no more than 30% in absolute value from that of the treatment group. We further require a minimum of two weeks of continuous data prior to the Yieldwatch token acquisition. The coefficient of 0.0497 is comparable to our baseline coefficient at the farm level reported in column (1) of Table 6. That is, a farm with a 100 points higher *Total Offered Yield* is associated with a 4.97% greater net increase in flows.

In column (2) of Table 8, we test whether more granular performance information and disclosure of hidden risks through WATCH token acquisition reduces reaching for yield behavior. The key coefficient of interest is the triple interaction between *Total Offered Yield* and two indicator variables that are equal to one if a farmer has ever owned Yieldwatch tokens or provided liquidity to the WATCH-BNB pool (*YieldWatch*), and after the Yieldwatch token acquisition (*Post*), respectively, and zero otherwise. The negative and statistically significant coefficient of -0.0290 suggests that, following the token acquisition, WATCH token holders with access to Yieldwatch.Pro exhibit a reduction in their reaching for yield propensity of about 58% (0.0290/0.0497) relative to investors without the WATCH tokens.

Our results are robust to controlling for farm times week fixed effects, allowing us to compare the reaching for yield propensity among WATCH token holders and non-holders while accounting for unobserved time-varying characteristics at the farm level. We also add interactions of treatment-control pair with week, farmer, and farm fixed effects to alleviate concerns of endogenous selection into farms and treatment-control pairs. Moreover, in columns (3) to (5), we impute the treatment effect using the three, five, and ten most closely matched control wallets (equally weighted).[8] Across specifications, the magnitude of the coefficient of interest barely changes and remains significant at the 1% level.

In Appendix Figure A.13 we plot the dynamics of the triple interaction coefficient between three quarters before and three quarters after the token acquisition. Before the acquisition, the estimated coefficients are statistically indistinguishable from zero, suggesting that there is a parallel trend in the reaching for yield propensity between wallets who do and do not hold WATCH tokens. Following the token acquisition, however, we observe that token holders significantly reduce their propensity

[7]To the extent that some investors are unsophisticated, the display of more information also changes investors' information sets. Enhanced information disclosure also reduces the perceived product complexity.

[8]Not all treated farmers have up to ten control wallets satisfying our matching criteria. We, therefore, implement equally-weighted regression to mitigate the impact of differential sizes in control groups. Our results are robust to non-weighted regressions.

to chase yields. Estimates become noisier for more distant periods following the access to more granular risk and performance information.

In columns (6) to (10) of Table 8, we further mitigate concerns that the token acquisition may be correlated with investor skill using a quasi-natural experiment. Specifically, we exploit the Yieldwatch initial farm offering (IFO) on March 4, 2021 that resulted in a random allocation of 8 Yieldwatch tokens to investors who offered below or above \$569.4 USD. In Appendix C.4, we provide details about the institutional design of the IFO and supportive evidence that the allocation of WATCH tokens around the threshold was as good as random. Importantly, we find no evidence of strategic bidding over the time of the auction and no differential bidding behavior by successful and unsuccessful bidders. Hence we can consider bidders just below and above the threshold as similar and examine their change in reaching for yield propensity after the token acquisition.

In columns (6) to (8), we examine differences in the reaching for yield propensity between successful and unsuccessful bidders for the Yieldwatch tokens in a narrow window of $+/-$ \$200 around the bidding threshold. That narrow band contains 148 treated (i.e., successful) and 86 control wallets. The coefficient on the triple interaction term is again negative and statistically significant at the 1% level, providing further evidence that better risk disclosure decreases the salience of prices in a relative sense, and reduces the propensity to reach-for-yield. In these regressions, we add farmer and farm fixed effects to absorb unobserved time-invariant heterogeneity. But our results do remain robust in the conservative specification reported in column (8) of Table 8, where we account for latent time-varying characteristics at the farm level using the inclusion of farm times week fixed effects.

In column (9), we extend the bandwidth to plus and minus \$250 around the allocation threshold and find that the coefficient remains significant and of similar magnitude. As we expand the bandwidth further to plus and minus \$300, the coefficient is no longer significant.

Taken together, our evidence consistently shows that yield-chasing behavior becomes less pronounced once investors access more complete information on their yield farming portfolios, specifically more detailed information on the determinants of returns that tend to be hidden and are associated with downside risks (e.g., impermanent loss). This result is consistent with the hypothesis that investors chase yield because they are salient thinkers (Bordalo, Gennaioli, and Shleifer, 2016). We show that investor's reliance on salient features of financial products in their decision-making can be reduced by the increased availability of information on other less-salient features through third-party information services.

6.3 Randomizing the information acquisition through airdrops

One concern with our analysis is that more sophisticated investors are more likely to know about Yieldwatch.Pro and are more likely to acquire Yieldwatch tokens. In that case, our results could be explained by unobserved differences in financial sophistication rather than salience shocks introduced by changes in displayed information about risks and historical investment performance that arguably reduce perceived product complexity.

Farmer fixed effects should absorb time-invariant differences in sophistication across investors. Moreover, the quasi-random Yieldwatch token allocation to ex-ante similar bidders in the Yieldwatch IFO should mitigate concerns of endogeneous adoption. We further address this concern by exploiting a natural experiment based on airdrops organized by APY.Vision. Airdrops are events in which APY.Vision provides a select group of users with APY.Vision NFT tokens granting access to premium tracking services. APY.Vision operates across multiple platforms and selects users randomly. PancakeSwap is not covered by APY.Vision, which operates on Ethereum, a different blockchain, and we manually collect all the airdrop announcements through X (formerly Twitter). We, therefore, focus our data collection efforts on wallets present in SushiSwap. See Appendix C.5 for details.

Access to the APY.Vision information service is enabled through the ownership of NFT tokens. APY.Vision randomly allocates these NFT tokens to liquidity providers in eligible liquidity pools. During our sample period, we identify 20 airdrops, in which 38 among all eligible wallets were randomly chosen to receive NFTs. We compare the reaching for yield propensity between the 38 investors with access to APY.Vision and 14,266 unsuccessful wallets in Sushiswap that were eligible to receive the NFT tokens. Since many treated wallets do not have yield farming history before the airdrops, we compare their cross-sectional differences in yield chasing propensity.

Table 9 reports results that are qualitatively similar to those obtained for the Yieldwatch experiment. First, the baseline magnitude of the coefficient in column (1), 0.0285 is reasonably similar to that reported in Table 7, even though the data come from a different decentralized market operating on a different blockchain. This is reassuring and supports external validity of our findings. Second, the interaction term between total offered yield and APY.Vision NFT token holder dummy variable is negative and statistically significant.

In column (3), we show that this result is robust to the more stringent specification with farm times week fixed effects. For columns (4) to (6), we select control groups based on the eligibility requirements and investment duration, since our results in Table 7 suggest that experience impacts reaching for yield. This analysis compares the behavior of unsuccessful farmers to those who happened to be lucky and received NFTs. We observe consistent results in the baseline specification in column (4), and in a more conservative specification in (5) using farm times week fixed effects. To further mitigate concerns of unbalanced control groups, we show consistent results using a weighted regression specification in column (6). Overall, these findings similarly suggest that a shock to investors' information access that reduces perceived complexity can significantly reduce investors' salience bias.

7 Conclusion

We provide the first characterization of yield farming, a decentralized financial service available to retail investors in the cryptocurrency ecosystem. Using a novel hand-collected dataset of all trade records in 262 yield farms listed on PancakeSwap, the largest automated market-maker operating on the Binance Smart Chain, we assess yield farming's return performance and document its associated risks.

Yield farms offer high yields that are saliently advertised as headline rates, while downside risks are hidden and not easily understood. Yield farming appears to be profitable, but risk-adjusted returns are significantly reduced after accounting for transaction fees, price impact, and investor mistakes. Investor flows are attracted to high yield farms but are insensitive to impermanent losses, a type of hidden downside risk. But, high yield farms systematically underperform due to large impermanent losses. Thus, we document reaching for yield behavior that results in negative risk-adjusted returns.

By means of two quasi-natural experiments designed by third-party information platforms, we study how information shocks that increase risk disclosure and arguably reduce complexity can affect yield-chasing behavior. We find consistent evidence that farmers' propensity to reach for yield becomes less pronounced once they are provided more detailed information on the performance of their portfolios. We also document evidence that investor learning and experience contributes to reducing their yield-chasing behavior over time.

Our results have important policy implications. First, our evidence emphasizes the need for better information disclosure, since it can mitigate investors' salience bias. Notably, the type of information matters, since our findings highlight the role of risk disclosure as opposed to price disclosure. In contrast, Frydman and Wang (2020) show that enhanced information about prices can increase the bias related to the disposition effect. Nonetheless, our results suggest that, even without regulatory mandates on information provision, market-based alternatives, such as third-party information platforms, can help assuage yield-chasing behavior and improve investment performance. Second, our evidence on investor mistakes and learning emphasizes the importance of financial education, especially for retail investors. Third, mandatory reductions in product complexity and proactive notifications can help overcome investors' inattention or inertia and, therefore, improve their performance.

References

Agarwal, Sumit, Itzhak Ben-David, and Yao Vincent, 2017, Systematic mistakes in the mortgage market and lack of financial sophistication, *Journal of Financial Economics* 123, 42–58.

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

Allen, Franklin, and Douglas Gale, 1994, (*Financial Innovation and Risk Sharing*).

Amromin, Gene, Jennifer Huang, Clemens Sialm, and Edward Zhong, 2018, Complex Mortgages, *Review of Finance* 22, 1975–2007.

Angeris, Guillermo, Hsien-Tang Kao, Rei Chiang, Charlie Noyes, and Tarun Chitra, 2021, An Analysis of Uniswap Markets, *Cryptoeconomic Systems* 1, published online.

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, 2023, The Impact of Derivatives on Spot Markets: Evidence from the Introduction of Bitcoin Futures Contracts, *Management Science* 69, 6417–7150.

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

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2012, Salience, *Quarterly Journal of Economics* 127, 1243–1285.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2013, Salience and consumer choice, *Journal of Political Economy* 121, 803–843.

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

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2022, Salience, *Annual Review of Economics* 14, 521–544.

Calvet, Laurent, Claire Célérier, Paolo Sodini, and Boris Vallée, 2023, Can Security Design Foster Household Risk-Taking?, *Journal of Finance* 78, 1917–1966.

Campbell, John, 2006, Household Finance, *Journal of Finance* 61, 1553–1604.

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

Carlin, Bruce, 2009, Obfuscation, Learning, and the Evolution of Investor Sophistication, *Journal of Financial Economics* 91, 278–287.

Carlin, Bruce, Shimon Kogan, and Richard Lowery, 2013, Trading Complex Assets, *Journal of Finance* 68, 1937–1960.

Carlin, Bruce, and Gustavo Manso, 2011, Obfuscation, Learning, and the Evolution of Investor Sophistication, *Review of Financial Studies* 24, 754–785.

Célérier, Claire, Gordon Liao, and Boris Vallée, 2022, The Price Effects of Inovative Security Design, *Working Paper Harvard University and University of Toronto*.

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, 2023, Reaching for Yield and Bond Returns, *Management Science* Forthcoming.

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

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

Cong, Lin William, Campbell R Harvey, Daniel Rabetti, and Zong-Yu Wu, 2022, An Anatomy of Crypto-Enabled Cybercrimes, *NBER Working Paper 30834*.

Cong, Lin William, Zhiheng He, and Ke Tang, 2022, Staking, Token Pricing, and Crypto Carry, *Working Paper*.

Cong, Lin William, Ke Tang, Yanxin Wang, and Xi Zhao, 2022, Inclusion and democratization through web3 and defi? initial evidence from the ethereum ecosystem, *Initial Evidence from the Ethereum Ecosystem (July 29, 2022)*.

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

Duffie, Darrell, and Ming Huang, 1995, FinancialMarket Innovation and Security Design: An Introductionï¿œ, *Journal of Economic Theory* 65, 1–42.

Efing, Matthias, 2020, Reaching for Yield in the ABS Market: Evidence from German Bank Investments, *Review of Finance* 24, 929–959.

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

Fang, Chuck, 2023, Liquidity Misallocation on Decentralized Exchanges, *SSRN Working Paper 4821293.*

Foley, Sean, Peter O'Neill, and Talis Putnins, 2022, Can Markets be Fully Automated? Evidence From an "Automated Market Maker", *Working Paper.*

Frydman, Cary, and Baolian Wang, 2020, The Impact of Salience on Investor Behavior: Evidence from a Natural Experiment, *The Journal of Finance* 75, 229–276.

Gabaix, Xavier, and David Laibson, 2006, Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets, *Quarterly Journal of Economics* 121, 505–540.

Gao, Pengjie, Allen Hu, Peter Kelly, Cameron Peng, and Ning Zhu, 2023, Asset Complexity and the Return Gap*, *Review of Finance* 28, 511–550.

Gensler, Gary, 2021, Remarks Before the Aspen Security Forum, Remarks by Chairman Gary Gensler at the Annual Aspen Security Forum, 3 August 2021.

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

Gomes, Francisco, Cameron Peng, Oksana Smirnova, and Ning Zhu, 2022, Reaching for Yield: Evidence from Households, *SSRN W0rking Paper 4283008.*

Griffin, John, Richard Lowery, and Alessio Saretto, 2014, Complex Securities and Underwriter Reputation: Do Reputable Underwriters Produce Better Securities?, *Review of Financial Studies* 27, 2872–2925.

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

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

Hasbrouck, Joel, Fahad Saleh, and Thomas Rivera, 2022, The Need for Fees at a DEX: How Increases in Fees Can Increase DEX Trading Volume, *Working Paper.*

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.

Inderst, Roman, and Marco Ottaviani, 2009, Misselling through Agents, *American Economic Review* 99, 883–908.

Inderst, Roman, and Marco Ottaviani, 2022, Excessive Competition for Headline Prices, *International Economic Review* 64, 883–908.

John, Kose, Leonid Kogan, and Fahad Saleh, 2023, Smart Contracts and Decentralized Finance, *Annual Review of Financial Economics* 15, 523–542.

Kacperczyk, Marcin T., and Marco Di Maggio, 2017, The Unintended Consequences of the Zero Lower Bound Policy, *Journal of Financial Economics* 123, 59–80.

Korevaar, Matthijs, 2023, Reaching for yield and the housing market: Evidence from 18th-century Amsterdam, *Journal of Financial Economics* 148, 273–296.

Kruppa, Miles, 2022, DeFi projects rife with hidden risks, global regulatory body warns, *Financial Times* 24 March 2022.

Lehar, Alfred, and Christine A. Parlour, 2024, Decentralized Exchanges, *Journal of Finance* Forthcoming.

Li, Thomas, Siddharth Naik, Andrew Papanicolaou, and Lorenzo Schönleber, 2024, Yield Farming for Liquidity Provision, *SSRN Working Paper 4422213*.

Liu, Yukun, Aleh Tsyvinski, and Xi Wu, 2022, Common Risk Factors in Cryptocurrency, *Journal of Finance* 77, 1133–1177.

Makarov, Igor, and Antoinette Schoar, 2022, Cryptocurrencies and decentralized finance (DeFi), *Brookings Papers on Economic Activity*.

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, *Financial Times* 2 July 2021.

Osipovich, Alexander, 2021, Crypto 'Yield Farmers' Chase High Returns, but Risk Losing It All, *Wall Street Journal* 17 July 2021.

Park, Andreas, 2023, The Conceptual Flaws of Constant Product Automated Market Making, *Management Science* 69, 6731–6751.

Sato, Yuki, 2014, Opacity in Financial Markets, *Review of Financial Studies* 27, 3502–3546.

Schwenkler, Gustavo, Aakash Shah, and Darren Yang, 2023, A Study Of The Reliability of Crypto Data Provision, *Available at SSRN 4673905*.

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.

Vokata, Petra, 2021, Engineering lemons, *Journal of Financial Economics* 142, 737–755.

Vokata, Petra, 2023, Do Investors Read the Fine Print? Salient Thinking and Security Design, *Ohio Sytate University Working Paper*.

Figure 1: Growing Popularity of Decentralized Finance

In this figure, we plot the total value locked (TVL, left axis), a measure of market capitalization, and the number of active platforms (right axis) in the market for decentralized finance. The solid blue line plots total value locked (TVL) in billions of dollars. The dashed red line illustrates the number of DeFi platforms whose TVL is over $1 million. We source historical TVL data from DeFiLlama API (https://api.llama.fi/v2/historicalChainTvl). To construct the number of active DeFi platforms, we first download each platform's TVL from DeFiLlama API on historical TVL for platforms (https://api.llama.fi/protocol/rainbow). Among the platforms, we drop "CEX" and "Liquidity Staking" following DeFiLlama's approach to construct a conservative measure of TVL. Each day, we count the number of remaining platforms whose TVLs are above $1 million. The figure starts on January 1, 2020 and ends on July 31, 2022. Source: https://defillama.com/.

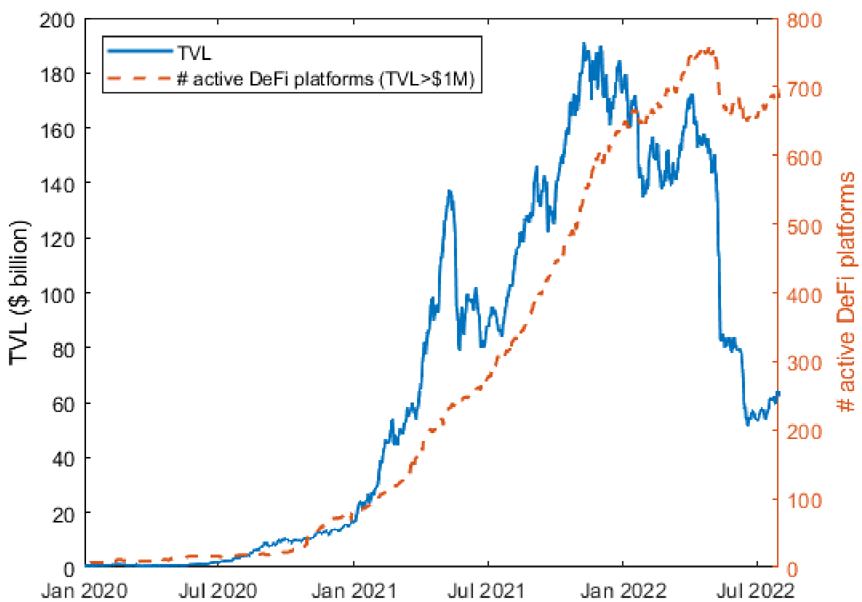


Figure 2: Average Gas Fee to Enter and Exit a Yield Farming Position

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 $1.49 ($1.96). For SushiSwap, the average cost to enter (exit) over all days is $117.75 ($178.10).

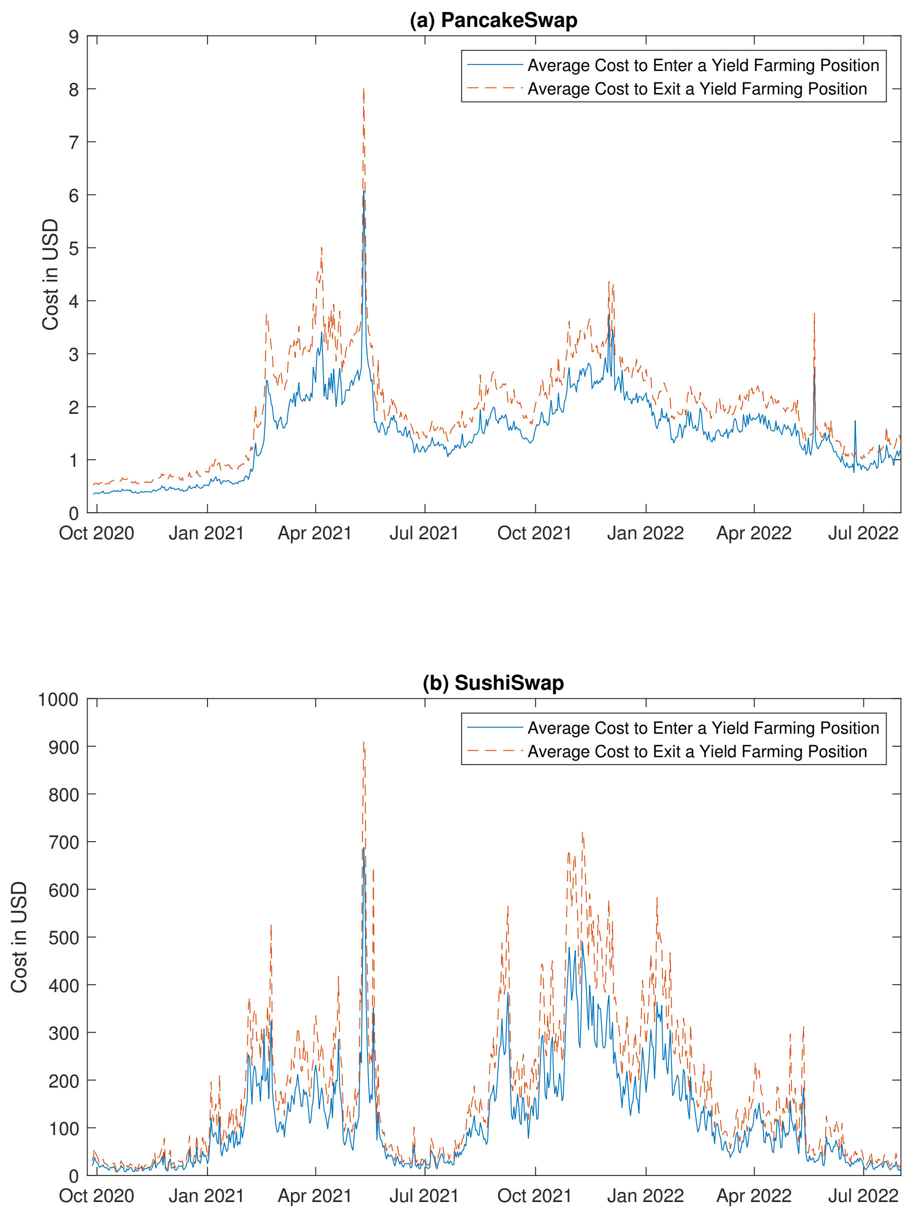


Figure 3: Heuristic Description of Yield Farming in PancakeSwap

This figure provides a heuristic description of yield farming in the decentralized exchange (DEX) PancakeSwap v2, which is built on the Binance Smart Chain (BSC). In PancakeSwap, investors face a menu of liquidity pools, each one being defined for a pair of cryptocurrencies. Our illustration showcases the USDT–ETH pool as an example. Investors can provide liquidity by "staking" a pair (x, y) of cryptocurrency tokens (in this example, USDT and ETH) in equal dollar amounts $(x \cdot P^{USDT} = y \cdot P^{ETH})$ into the liquidity pool, thereby making these tokens available for USDT–ETH trading by third-party traders. These must pay a trading fee for buying and selling USDT vs. ETH equal to 0.25% of trading volume. Of the 0.25% trading fee, 0.17% is paid to liquidity providers as compensation for their liquidity provision. The other 0.08% is passed on to the Treasury of PancakeSwap's main staking contract and partially used for burning (i.e., buy back and destroy) CAKE tokens, the native governance token of PancakeSwap. The main staking contract issues CAKE tokens on a continuous basis with each block creation in BSC. The trading fees are paid in the currency of the liquidity pool, i.e., USDT vs. ETH. As a liquidity provider, an investor faces buy-and-hold price risk from the price evolution of USDT and ETH as well as downside risk arising from the impermanent loss function, defined by the constant product trading rule of the automated market maker (AMM). The liquidity provision is certified by a liquidity token (i.e., the LP token), which can be staked into a USDT–ETH main staking contract (yield farm) specific to the USDT-ETH currency pair. The passive income in the yield farm is earned in CAKE. The number of CAKE tokens distributed across yield farms depends on a farm multiplier that we describe in Section 2. This farm multiplier may change over time following a collective vote by all owners of CAKE tokens.

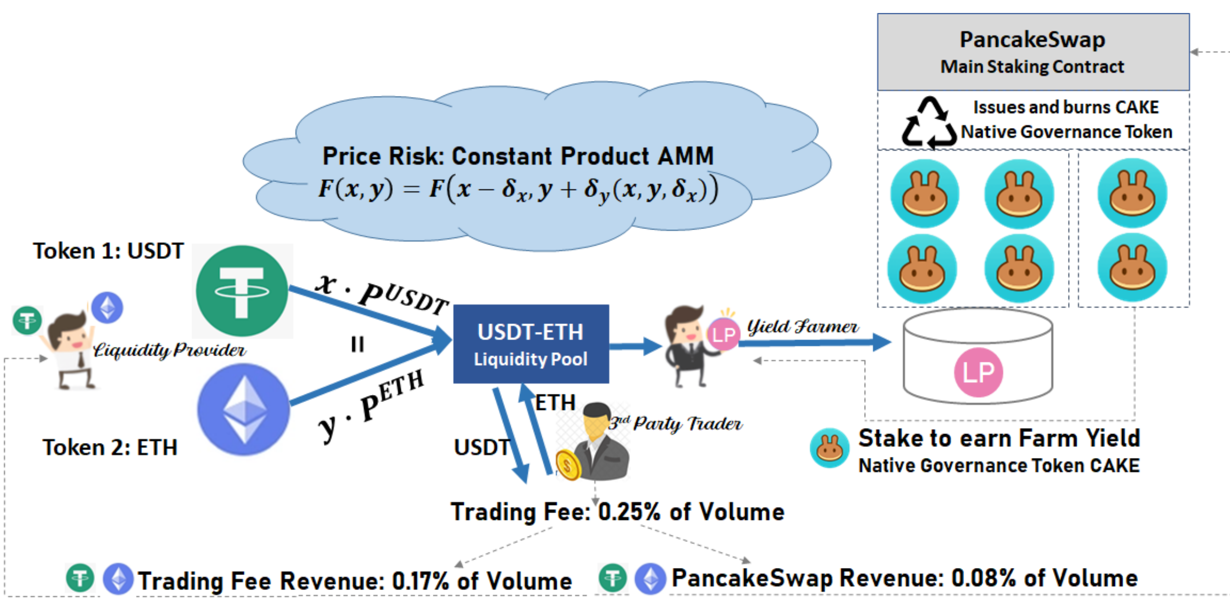


Figure 4: Yield Farm Activity

In Panel (a) of this figure, we plot the number of active farms and Total Value Locked (TVL) in $million at a weekly frequency during our sample period. On the right axis, we provide the time series of active farms. We define active farms as those whose yield multipliers are larger than 0, implying that investors who stake LP tokens in these farms receive non-negative yields. On the left axis, we plot TVL of active farms, or the amount of LP tokens deposited for yield farming. The vertical axis is in millions of USD. In Panel (b) of this figure, we illustrate the Google search intensity for the word, "PancakeSwap," and the number of active farmers in PancakeSwap. We download the Google search intensity for the word, "PancakeSwap," and calculate the monthly average search intensity. Then, we normalize it by the maximum monthly average search intensity so that the index is 100 at its maximum. The dotted blue line (left axis) plots the normalized monthly average of the search intensity. Google search data are available at https://trends.google.com/trends/explore?q=PancakeSwap. The dashed red line (right axis) plots the number of active farmers, where an active farmer is defined to be an investor whose balance in yield farms is positive. The red dotted vertical line corresponds to March 1, 2021. The figures start on September 23, 2020 and end on July 31, 2022.

(a)

(b)

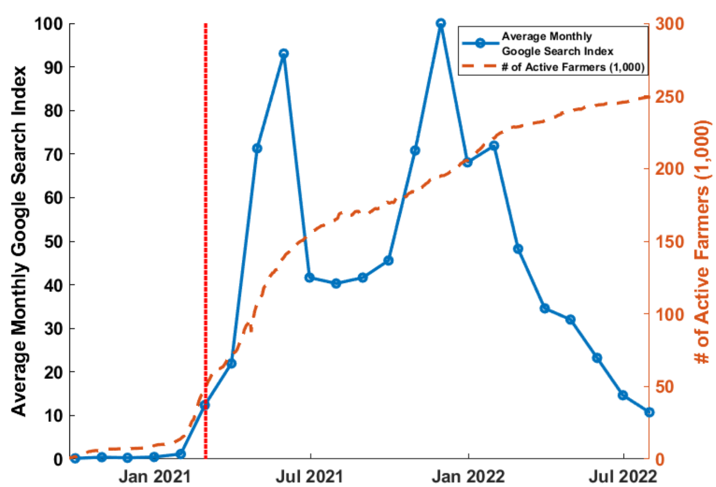


Figure 5: Total Offered Farm Yields Displayed to Investors

In this figure, we plot the annualized total offered farm yields displayed to yield farmers. Total offered yield, referred to as the annual percentage return (APR), is the sum of the offered farm yield (Equation (6)) and the trading fee yield estimated using the previous day's trading volume $V_{t-1,t}$.

$$y_t^{Total} = \underbrace{y_t^{Farm}}_{\text{offered farm yield}} + \underbrace{365 \cdot c \cdot V_{t-1,t}/L_t}_{\text{trading fee yield}}, \tag{12}$$

where c denotes the constant trading fee and L_t the pool liquidity. We provide the historical annualized total offered farm yields (in %) between March 1, 2021, and July 31, 2022. The solid blue line indicates the median annualized total offered farm yield. Dark and light shaded areas represent the interquartile range, as well as the 10th and 90th percentiles of the yield distribution, respectively.

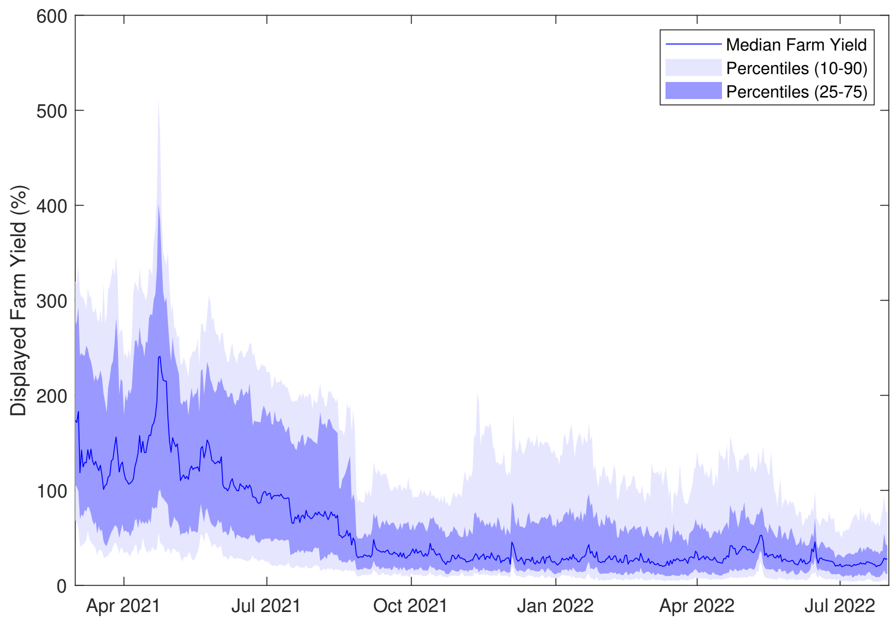


Figure 6: Migration of PancakeSwap Platforms

In this figure, we show the amount of outstanding liquidity in obsolete liquidity pools after two technical updates in the PancakeSwap platform. In Panel (a), we plot the total value locked in liquidity pools and their associated yield farms in PancakeSwap v1 with their equivalent counterpart yield farms available in PancakeSwap v2. On April 24, 2021, farms corresponding to liquidity pools in PancakeSwap v1 stopped providing farm yields. PancakeSwap encouraged farmers to move their liquidity to the corresponding counterpart farms available in PancakeSwap v2 so that the existing yield farmers could continue to earn farm yields. The solid blue line in Panel (a) indicates the total value locked of unmigrated assets that remained in the liquidity pools associated with PancakeSwap v1. In Panel (b), we examine the outstanding liquidity staked in the old PancakeSwap staking contract, following the contract's upgrade from v1 to v2 on April 20, 2022. Upon this migration, LP tokens staked in the old staking contract ceased to be eligible for earning yields. PancakeSwap advertised through Twitter and other channels that users should unstake from the v1 contract and re-stake in the new v2 contract.

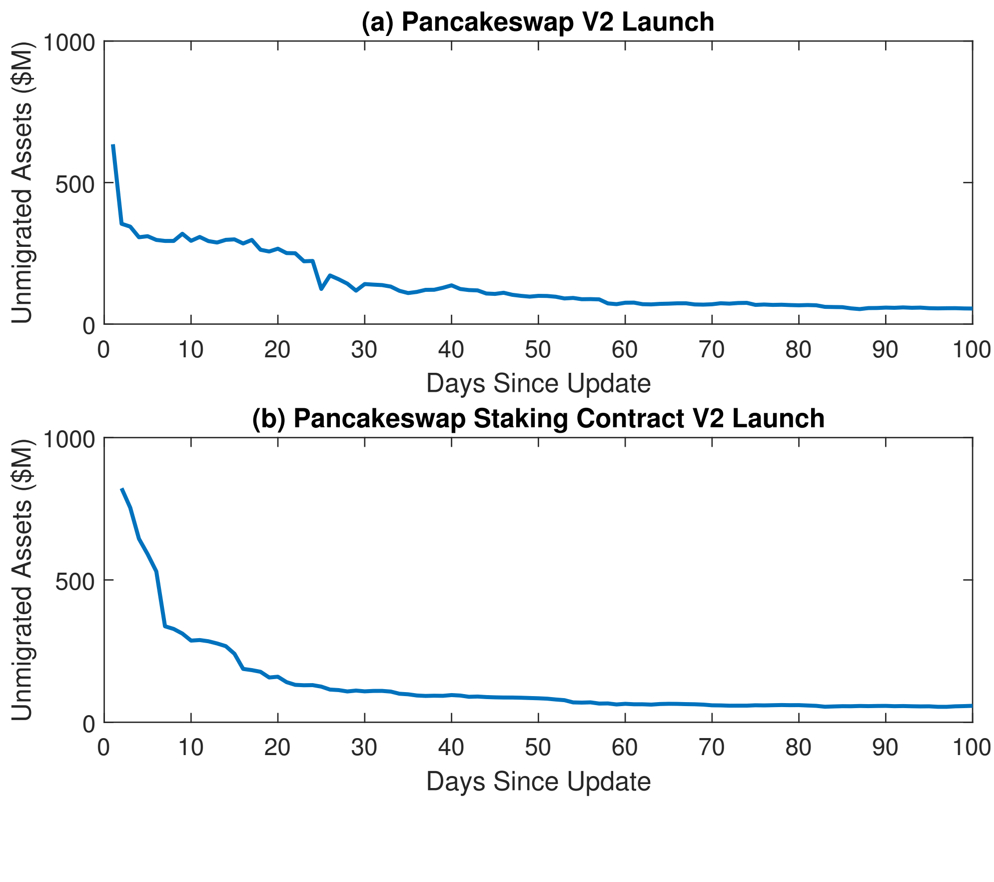


Figure 7: Staking Ratio of LP Tokens

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 staking ratio. Dark and light shaded areas represent the interquartile range, as well as the 10th and 90th percentiles of the yield farm distribution, respectively.

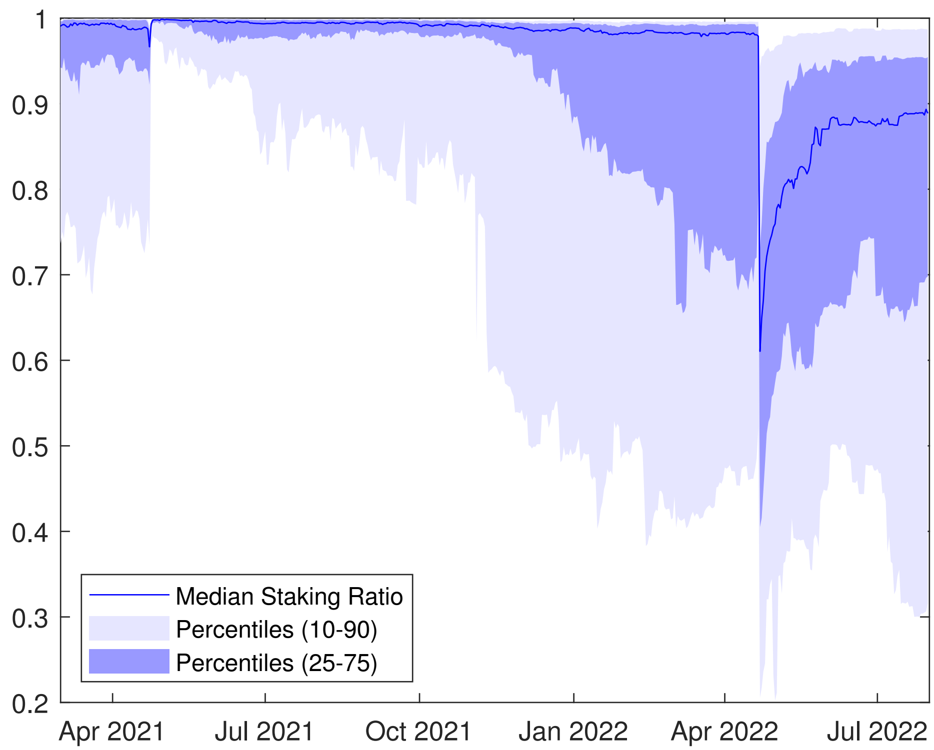


Figure 8: Yield Farming Return Decomposition

In this figure, we plot each component of daily value-weighted returns across yield farms, by quintiles based on the magnitude of their in-sample offered yield. Every day, we compute the daily capital gain, impermanent loss, trading fee, and realized yield for all listed farms. Then, we take the average of each component across farms in each quintile using the size of each farm as weights. In Panels (a) to (d), the blue bars illustrate the average daily realized yield, trading fee, capital gain, and impermanent loss. The red error bars plot their associated 95% confidence intervals. The mean of each component is displayed above their respective bars.

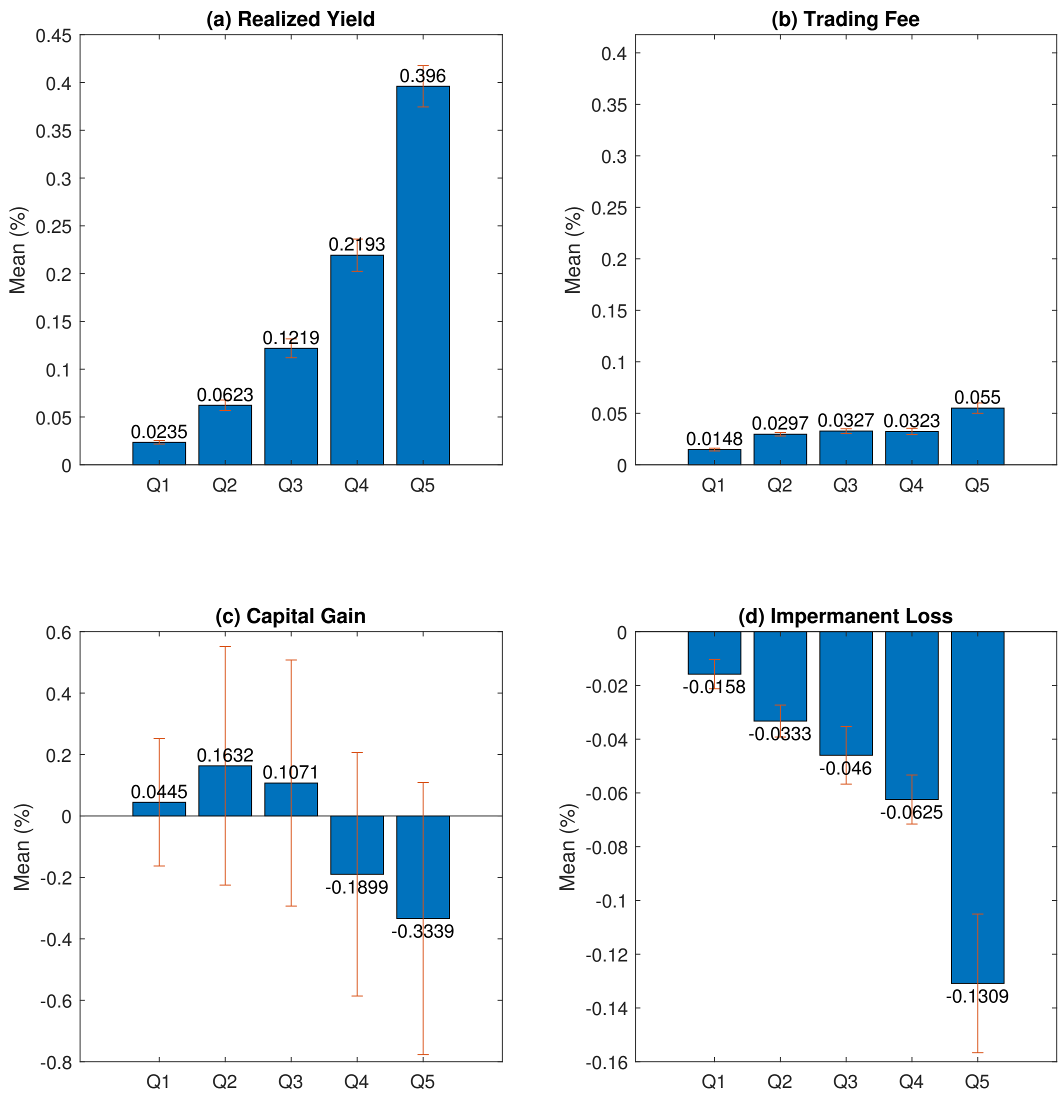


Figure 9: Risk-Adjusted Returns from Yield Farming

In this figure, we plot average risk-adjusted returns (i.e., alphas) and their associated 90% & 95% confidence intervals for different trading strategies. In Panel (a), we compare the performance of yield farming to that of liquidity mining without considering trading frictions. On each day, we sort farms into quintiles based on their in-sample total offered farm yields (APRs) displayed to investors. In each quintile, we form value-weighted portfolios by using the size of the liquidity pools as weights. A yield farming strategy is a strategy in which investors not only earn trading fee revenue but also farm yields, whereas investors that restrict themselves to liquidity mining can only earn trading fee revenue. We estimate alphas from a three factor model based on the work of Liu, Tsyvinski, and Wu (2022) and also account for the performance of BNB. The circle (square) and the associated bar display weekly alphas and their 95% confidence intervals for yield farming (liquidity mining) without considering frictions. In Panel (b), we follow a similar procedure but provide alphas for yield farming strategies without trading frictions, yield farming strategies with frictions including gas fees, trading fees, and price impact, and yield farming strategies considering not only the frictions but also investor mistakes.

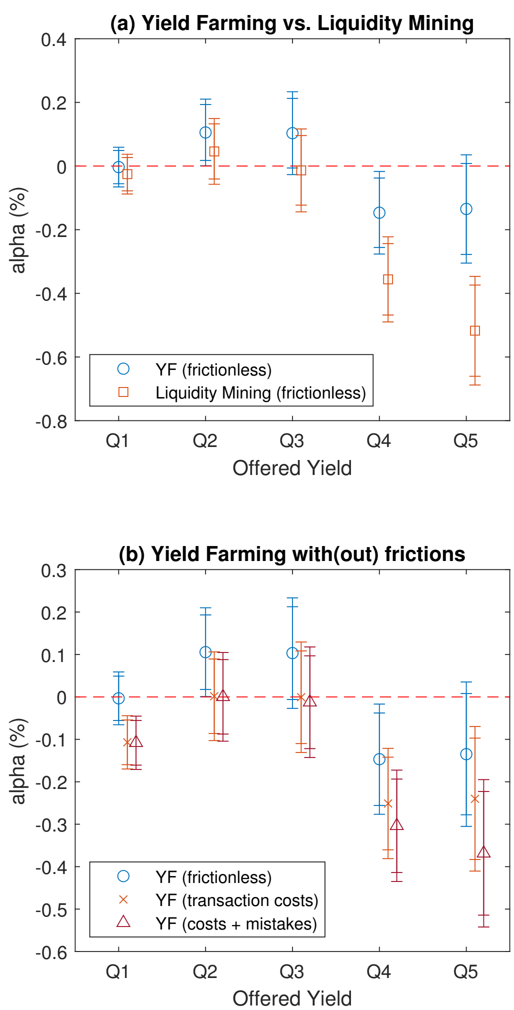
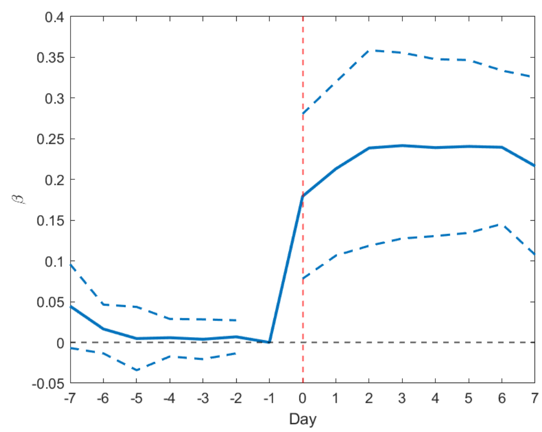


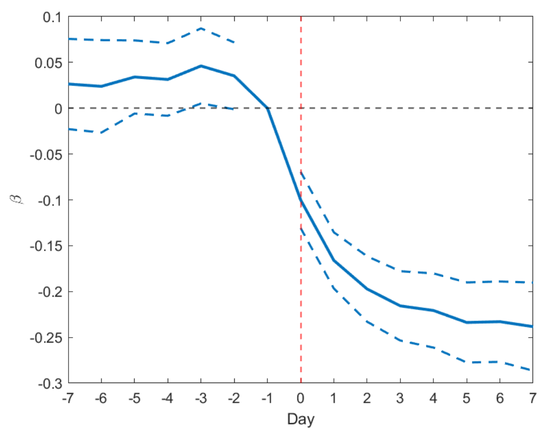
Figure 10: The Impact of Farm Multiplier Changes on Farm Flows

In this figure, we illustrate, using an event study, how changes in the CAKE allocation multipliers, $\Delta m_{i,t}$, affect cumulative flows to the farm. In Panels (a) and (b), we measure cumulative flows net of any price growth from one day before the event date $(t-1)$ to date $t+h$ by measuring the change of # LP tokens. In Panels (c) and (d), we measure the growth rate of the dollar value of the pool. We are interested in changes in cumulative flows and dollar value of the pools that are driven by active decisions of PancakeSwap platform owners while there is no significant change to the aggregate multiplier, that is $\Delta m_{i,t} \neq 0$ with $|\frac{\Delta M_t}{M_t}| \leq 0.15$. We identify 511 such cases, among which 50 cases are associated with an increase in $m_{i,t}$, and 461 cases are associated with a decrease in $m_{i,t}$. We then compare the change in the outcome variables of the treated farms relative to those of the non-treated farms (all farms that did not experience a multiplier change). Specifically, we plot the difference-in-differences coefficients β_k (and their 95% confidence intervals) from a regression $y_{i,t+h} = \alpha + \sum_{k=-7,k\neq-1}^{k=7} \beta_k I_k \times Treatment_i + Event \times FarmFE + DayFE + \varepsilon_{i,t+h}$, where $y_{i,t+h}$ is defined as either $log(\frac{outstanding\ LP\ tokens_{i,t+h}}{outstanding\ LP\ tokens_{i,t-1}})$ or $log(\frac{\$\ of\ pool_{i,t+h}}{\$\ of\ pool_{i,t-1}})$. We cluster the standard errors at the farm and date levels. Panels (a) and (c) are event studies for $\Delta m_{i,t} > 0$ while Panels (b) and (d) are event studies for $\Delta m_{i,t} < 0$.

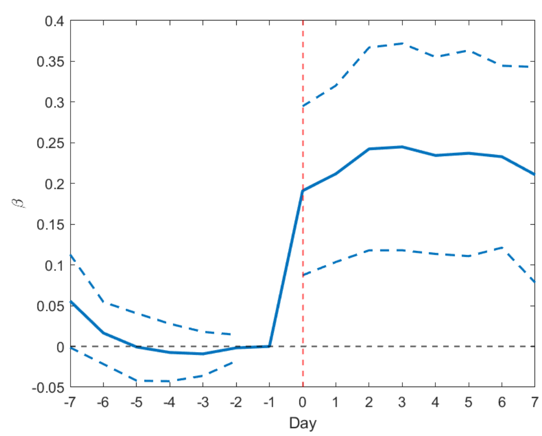
(a) $\Delta m_{i,t} > 0$



(b) $\Delta m_{i,t} < 0$



(c) $\Delta m_{i,t} > 0$



(d) $\Delta m_{i,t} < 0$

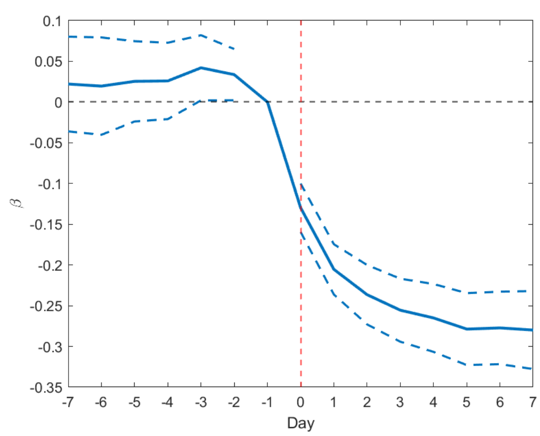


Table 1: Snapshot of Yield Farms in PancakeSwap

In this table, we report information about the 10 largest farms in PancakeSwap in terms of total value locked (TVL, Panel A) or total offered farm yield (Panel B) on July 31, 2022. For each farm, defined by a unique cryptocurrency pair, we provide information on the start date of a farm, the annualized total offered farm yield (APR, in %), and total value locked (TVL, in $ million). Panel A lists the 10 largest famrs in terms of TVL. Panel B lists the 10 largest farms in terms of total offered farm yield (APR). The total offered farm yield is the sum of the offered yield and the trading fee yield estimated using the previous day's trading volume.

Panel A: By TVL

Farm Rank	**Cryptocurrency Pair**	**TVL ($ million)**	**Total Offered Yield (%)**
1	USDT-BUSD	$178.28M	2.62%
2	WBNB-BUSD	$168.35M	14.25%
3	Cake-WBNB	$168.18M	24.30%
4	USDT-WBNB	$158.34M	15.10%
5	USDC-BUSD	$109.69M	1.18 %
6	USDT-USDC	$53.99M	2.17%
7	ETH-WBNB	$53.46M	7.17%
8	BTCB-WBNB	$45.24M	7.38%
9	BTCB-BUSD	$43.59M	9.79%
10	TUSD-BUSD	$36.67M	0.30%
...	...	...	...
86	GMI-WBNB	$0.12M	81.81%

Panel B: By Total Offered Yield

Farm Rank	**Cryptocurrency Pairs**	**TVL ($ million)**	**Total Offered Yield (%)**
1	BTCST-WBNB	$1.72M	357.92%
2	OLE-BUSD	$1.20M	138.87%
3	TRIVIA-WBNB	$0.89M	127.02%
4	HIGH-BUSD	$1.18M	99.87%
5	XWG-USDC	$0.69M	90.00%
6	RPG-BUSD	$1.11M	87.33%
7	IDIA-BUSD	$0.17M	83.23%
8	GMI-WBNB	$0.12M	81.81%
9	FINA-BUSD	$0.40M	76.70%
10	BCOIN-WBNB	$0.24M	71.89%
...	...	...	...
86	TUSD-BUSD	$36.67M	0.30%

Table 2: Determinants of Farm Yields driven by Platform Governance

In this table, we study the determinants of farm yield changes associated with active platform governance ($\Delta y^m_{i,t+1}$), i.e., the component of farm yield changes associated with changes in the farm yield multiplier m. This is computed as the product between the current yield level and the percentage change of the yield multiplier, i.e., $\Delta y^m_{i,t+1} = y_{i,t} \times \frac{\Delta m_{i,t+1}}{m_{i,t}}$. In columns (1) and (2), the dependent variable is the change in yield that is driven by platform governance. In columns (3) and (4), the dependent variable is *Delisting*$_{t+1}$, an indicator variable equal to one if a farm is delisted on the subsequent day and zero otherwise. Independent variables include *Capital Gain*, *Impermanent Loss*, *Trading Fee*, *Realized Yield* over the last 7 days, and *Log(Liquidity)*, which is the logarithm of the dollar value of aggregate liquidity in a pool. Standard errors are clustered at the farm level. *,**, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)
	$\Delta y^m_{i,t+1}$		Delisting$_{t+1}$	
Capital Gain$_{t-7,t}$	0.0067	0.0024	0.0050*	-0.0033
	(0.0060)	(0.0109)	(0.0027)	(0.0046)
Impermanent Loss$_{t-7,t}$	0.0877	0.0753	-0.0031	-0.0163
	(0.0607)	(0.0649)	(0.0114)	(0.0147)
Trading Fee$_{t-7,t}$	0.3282***	0.2475**	-0.2013***	-0.1552**
	(0.1064)	(0.1043)	(0.0665)	(0.0714)
Realized Yield$_{t-7,t}$	-0.5943***	-0.7692***	0.0546	0.0247
	(0.0870)	(0.1190)	(0.0340)	(0.0402)
log(Liquidity$_t$)	-0.0009**	-0.0012*	-0.0019***	-0.0026***
	(0.0004)	(0.0006)	(0.0003)	(0.0003)
log(Farm Age$_t$)	-0.0024***	-0.0042***	0.0003	0.0018***
	(0.0007)	(0.0007)	(0.0004)	(0.0004)
Day FE	No	Yes	No	Yes
N	48,770	48,770	49,048	49,048
adj. R^2	0.003	0.046	0.003	0.088

Table 3: Yield Farming Behavior

In this table, we report statistics that describe the behavior of yield farmers. The presented statistics are all farmer-level variables. In Panel A, we present aggregate summary statistics. *No. Farms* is the number of farms in which a yield farmer invests. *Investment Size* is the dollar value of LP tokens. *Holding Period (Days)* is the number of days for which a farmer keeps a yield farming position on average. *Offered Yield* is the time-weighted average of the total offered yield at the beginning of the holding period. The total offered farm yield is the sum of the offered yield and the trading fee yield estimated using the previous day's trading volume. *Daily Return* is the time-weighted average of the annualized holding period returns for each user. *Staking Ratio* is the average of the staking ratios of farms in which a farmer invested where the staking ratio of a farm is the average daily staking ratio during a farmer's holding period. In Panel B, we separate yield farmers into quintiles by *Investment Size*. We restrict our analysis to farmers whose investment duration is greater than or equal to one day.

Panel A: Yield Farmers

Variable	**Average**	**Inv. Size Weighted Avg.**	**SD**	**p25**	**p50**	**p75**	**OBS**
No. Farms	2.6363	4.7687	3.9766	1.0000	1.0000	3.0000	439,639
Investment Size ($)	7,732.14		231,923.65	40.00	179.66	869.44	439,639
Holding Period (Days)	30.9191	7.7183	64.3387	0.7086	3.4648	24.7954	439,639
Offered Yield	1.1002	0.5976	1.0534	0.4013	0.6995	1.4569	439,639
Daily Return	0.0011	0.0011	0.0569	-0.0035	-0.0005	0.0049	439,639
Staking Ratio	0.8422	0.9745	0.3385	0.9790	0.9992	0.9999	439,639

Panel B: Yield Farmers by Investment Size

	No. Farms	**Investment Size($)**	**Holding Period(Days)**	**Offered Yield**	**Daily Return**	**Staking Ratio**	**OBS**
Quintile 1							
Average	1.7430	10.96	61.3089	0.9428	0.0000	0.6211	87,928
SD	(1.6136)	(7.9)	(92.4108)	(1.0317)	(0.0267)	(0.4580)	
Quintile 2							
Average	1.8986	59.08	39.3546	1.0461	0.0006	0.8144	87,928
SD	(2.0561)	(21.47)	(70.5432)	(1.0567)	(0.0481)	(0.3581)	
Quintile 3							
Average	2.3041	187.72	26.5402	1.1382	0.0011	0.8821	87,927
SD	(2.896)	(62.03)	(56.2208)	(1.0574)	(0.0286)	(0.292)	
Quintile 4							
Average	2.9168	665.36	17.5689	1.2169	0.0020	0.9306	87,928
SD	(3.8799)	(259.88)	(41.4067)	(1.0746)	(0.1083)	(0.2234)	
Quintile 5							
Average	4.3190	37,737.51	9.8229	1.1571	0.0018	0.9629	87,928
SD	(6.6658)	(517,512.29)	(25.8638)	(1.024)	(0.0246)	(0.1613)	

Table 4: Yield Farming Performance

This table reports the summary statistics for daily 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 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, 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 (2022) augmented with the return of BNB, the native token of BSC, and the t-statistic for alpha from the three-factor+BNB regressions. The sample period is March 1, 2021 to July 31, 2022.

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	0.0015	0.0360	-0.0144	0.0023	0.0167	-0.3445	13.5692	-0.1638	0.0405	0.0003	1.0427	518
Buy and Hold (Capital Gains)	0.0007	0.0358	-0.0150	0.0018	0.0162	-0.2925	13.1318	-0.1681	0.0209	-0.0004	-1.1519	518
Liquidity Mining	0.0007	0.0358	-0.0148	0.0019	0.0162	-0.3959	13.5994	-0.1675	0.0197	-0.0004	-1.2605	518
Benchmark Strategy												
Crypto Market Return	-0.0002	0.0439	-0.0218	0.0049	0.0239	-0.7872	8.6170	-0.1026	-0.0052	0.0000	0.0000	518
BTC	-0.0005	0.0383	-0.0219	-0.0005	0.0200	-0.0654	4.7533	-0.0474	-0.0140	-0.0012	-1.2846	518
ETH	0.0016	0.0509	-0.0283	0.0023	0.0303	-0.0305	5.9682	-0.0522	0.0321	0.0007	0.7572	518
BNB	0.0021	0.0539	-0.0246	0.0013	0.0303	-0.0093	8.9147	-0.1232	0.0380	0.0000	0.0000	518
S&P 500 Index	0.0003	0.0116	-0.0059	NaN	0.0075	-0.3856	4.0643	-0.0157	0.0300	0.0004	0.7642	358
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	0.0105	0.0911	-0.0343	0.0076	0.0500	-0.4847	7.4880	0.1288	0.1147	0.0033	1.5869	74
Buy and Hold (Capital Gains)	0.0056	0.0904	-0.0380	0.0029	0.0470	-0.5332	7.6623	0.1158	0.0616	-0.0017	-0.8035	74
Liquidity Mining	0.0053	0.0898	-0.0372	0.0038	0.0473	-0.6369	7.8369	0.1099	0.0586	-0.0019	-0.8713	74
Benchmark Strategy												
Crypto Market Return	-0.0001	0.1110	-0.0764	0.0003	0.0800	-0.6809	4.6258	0.0948	-0.0005	0.0000	0.0000	74
BTC	-0.0047	0.0906	-0.0682	-0.0045	0.0479	-0.3788	2.9578	0.1803	-0.0516	-0.0034	-0.5156	74
ETH	0.0109	0.1298	-0.0771	-0.0032	0.0985	-0.1422	3.5841	0.1624	0.0840	0.0114	1.4707	74
BNB	0.0144	0.1402	-0.0636	0.0145	0.0829	-0.0877	6.3156	0.0606	0.1029	0.0000	0.0000	74
S&P 500 Index	0.0016	0.0239	-0.0131	0.0038	0.0155	0.1871	3.9293	-0.0739	0.0686	0.0019	0.8090	74

Table 5: Determinants of Yield Farmers' Return Performance

In this table, we study the determinants of the risk-adjusted return performance at the farmer level. The dependent variable, *Avg, Daily Ret. (w/o Frictions)*, is the time-weighted average daily holding period return for each farmer without considering trading frictions such as trading fees, gas fees, and price impact. *Avg. Daily Ret. (Frictions)* is the time-weighted average daily holding period return for each farmer considering the trading frictions. *Avg. Total Offered Yield* is the time-weighted average of the total offered farm yield displayed to investors (APRs) at the beginning of each holding period. The total offered farm yield is the sum of the offered yield and the trading fee yield estimated using the previous day's trading volume. *log(Avg. # of monthly Rebalancings)* is the average number of rebalancings in a month. *# of Farms* is the number of unique farms to which an investor provides liquidity. *Avg. Size of Investment* is the time-weighted average of the USD value at the beginning of an investor's holding period. *Avg. Staking Ratio* is the time-weighted average of the staking ratio of a farmer. We restrict our analysis to farmers whose investment duration is greater than or equal to one day. We include fixed effects for the entry month of the yield farming strategy (*Start Month*), the exit month of a yield farming investment (*End Month*), or the interaction of both. *Avg. Size of Investment*, *Avg. Daily Ret. (w/o Frictions)* and *Avg. Daily Ret. (Frictions)* are winsorized at the 0.25% and 99.75% levels. Standard errors are clustered at the first month when a farmer participated in yield farming. *,**, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Avg. Daily Ret. (w/o Frictions)			Avg. Daily Ret. (Frictions)		
Avg. Total Offered Yield	-0.0013**	-0.0015**	-0.0016***	-0.0032***	-0.0023**	-0.0024**
	(0.0005)	(0.0005)	(0.0005)	(0.0009)	(0.0009)	(0.0008)
# of Farms		-0.0000	-0.0000		0.0011***	0.0008***
		(0.0001)	(0.0001)		(0.0003)	(0.0002)
Avg. Size of Investment ($M)		-0.0231*	-0.0220*		0.3022***	0.3016***
		(0.0115)	(0.0113)		(0.0322)	(0.0309)
Avg. Size of Investment^2		0.0621*	0.0573*		-0.8669***	-0.8695***
		(0.0308)	(0.0306)		(0.0868)	(0.0837)
log(Avg. # of monthly Rebalancings)		0.0002	0.0004		-0.0061***	-0.0051***
		(0.0002)	(0.0003)		(0.0007)	(0.0006)
Avg. Staking Ratio		0.0025**	0.0023*		0.0204***	0.0193***
		(0.0010)	(0.0011)		(0.0014)	(0.0016)
Start Month	Yes	Yes	No	Yes	Yes	No
End Month	Yes	Yes	No	Yes	Yes	No
Start x End Month	No	No	Yes	No	No	Yes
N	439,639	439,639	439,639	439,639	439,639	439,639
adj. R^2	0.083	0.085	0.107	0.053	0.118	0.139

Table 6: Aggregate Farm Yields and Investor Flows

In this table, we report evidence on the relation between aggregate investor flows and total offered farm yields (APRs). We regress future farm *Flow*, measured over the next 7 days (a week), on *Total Offered Farm Yield*, past *Return* on yield farming, *Capital Gain*, *Impermanent Loss*, *Trading Fee Revenue*, and *Realized Yield* over the last 7 days, including control variables consisting of *Past flow*, *Log(Size of Liquidity Pool)*, *Farm age*. The total offered farm yield is the sum of the offered yield and the trading fee yield estimated using the previous day's trading volume. The sample period is March 1, 2021 to July 31, 2022. *Flow* and *Past flow* are winsorized at the 0.25% and 99.75% levels. Standard errors are clustered at the farm level. *,**, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)
		$Flow_{t,t+7}$	
Total Offered Yield$_t$	0.0635***	0.0638***	
	(0.0119)	(0.0119)	
Return$_{t-7,t}$		-0.0132	
		(0.0211)	
Capital Gain$_{t-7,t}$			-0.0169
			(0.0242)
Impermanent Loss$_{t-7,t}$			0.1019
			(0.1370)
Trading Fee$_{t-7,t}$			4.0200***
			(1.4435)
Realized Yield$_{t-7,t}$			2.4508***
			(0.4983)
Controls	Yes	Yes	Yes
Farm FE	Yes	Yes	Yes
Week FE	Yes	Yes	Yes
N	6,538	6,538	6,538
adj. R^2	0.148	0.148	0.142

Table 7: The Role of Learning and Experience in Reaching for Yield

In this table, we examine reaching for yield through the relation between *Flow* at the farmer level, measured over the next 7 days (a week) and *Total Offered Yield*, i.e., the total offered farm yield (APR). The total offered farm yield is the sum of the offered yield and the trading fee yield estimated using the previous day's trading volume. We further examine the role of learning in explaining the relation between flows by farmers and Offered Yield using the size of yield farming portfolio, investor experience defined as the number of days invested in yield farming, and the number of invested farms. *High Size* is an indicator variable equal to one if an investor's dollar value of the yield farming portfolio is greater than the 75th percentile of the size distribution and zero otherwise. *High Experience (days)* is an indicator variable equal to one if an investor's number of days elapsed since the start of the yield farming investment is greater than the 75th percentile of the distribution and zero otherwise. *High # Farms* is an indicator variable equal to one if the number of farms to which an investor has provided liquidity is greater than the 75th percentile of the distribution and zero otherwise. *Flow* is winsorized at the 0.25% and 99.75% levels. Standard errors are double clustered at the investor and week level.*,**, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				$Flow_{t,t+7}$			
Total Offered Yield	0.0255***	0.0278***		0.0254***		0.0303***	
	(0.0074)	(0.0065)		(0.0075)		(0.0067)	
High Size		-0.0713***	-0.0714***				
		(0.0049)	(0.0046)				
Total Offered Yield× High Size		-0.0029	-0.0032				
		(0.0033)	(0.0023)				
High Experience (days)				0.0219***	0.0177***		
				(0.0024)	(0.0017)		
Total Offered Yield× High Exp. (days)				-0.0091***	-0.0050**		
				(0.0033)	(0.0020)		
High # Farms						-0.0076*	-0.0045
						(0.0045)	(0.0042)
Total Offered Yield× High # Farms						-0.0123***	-0.0117***
						(0.0035)	(0.0028)
Farmer FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Farm FE	Yes	Yes	No	Yes	No	Yes	No
Week FE	Yes	Yes	No	Yes	No	Yes	No
Farm x Week FE	No	No	Yes	No	Yes	No	Yes
N	9,705,043	9,705,043	9,705,043	9,705,043	9,705,043	9,705,043	9,705,043
adj. R^2	0.199	0.201	0.229	0.199	0.227	0.199	0.227

Table 8: Yieldwatch Experiment: Information Display and Reaching for Yield

In this table, we investigate whether information on hidden risks and past performance displayed at YieldWatch.net impacts farmers' reaching for yield propensity using an event study analysis. $Flow_{t,t+7}$ defines investors' farm flows over the subsequent 7 days. *Yieldwatch* is an indicator variable equal to one if an investor/wallet holds the Yieldwatch token or provides liquidity to the Watch-BNB liquidity pool, and zero otherwise. *Post* is an indicator variable equal to one after the acquisition of Yieldwatch tokens and zero otherwise, defined in event time (i.e., one for positive event times and zero otherwise). *Total Offered Yield* is the total offered farm yield of a farm displayed to investors (APR). The total offered farm yield is the sum of the offered yield and the trading fee yield estimated using the previous day's trading volume. Investor controls include the natural logarithm of the dollar value of the yield farming portfolio of an investor and the natural logarithm of one plus the number of days since the starting date of the yield farming investment. In columns (1) to (5), we compare the Yieldwatch holders (treatment) to matched control wallets conditional on (i) data availability within two quarters before the Yieldwatch acquisition; (ii) a minimum of 2 weeks of data prior to the Yieldwatch acquisition; (iii) similar farm composition over the past 180 days (number and types of farms); (iv) total value locked one day before the event date deviating no more than 30% in absolute value from that of the treatment group. We add interactions of treatment-control pair (T-C Pair) and week, farmer, and farm fixed effects. In columns (6) to (10), we exploit the Yieldwatch initial farm offering (IFO) on March 4, 2021 that resulted in a random allocation of 8 Yieldwatch tokens to investors who offered below or above \$569.4 USD. We examine windows of +/- \$200, \$250 and \$300 USD around the threshold. *Flow* is winsorized at the 0.25% and 99.75% levels. Standard errors are double clustered at the investor and week level. *,**, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
						$Flow_{t,t+7}$				
Total Offered Yield	0.0497***					0.0537***	0.0282			
	(0.0095)					(0.0148)	(0.0282)			
YieldWatch×Total Offered Yield		0.0190**	0.0222***	0.0246***	0.0241***		0.0319	0.0620*	0.0531*	0.0067
		(0.0091)	(0.0071)	(0.0070)	(0.0069)		(0.0309)	(0.0323)	(0.0304)	(0.0307)
Post×Total Offered Yield		0.0694***	0.0784***	0.0745***	0.0674***		0.0533**			
		(0.0156)	(0.0121)	(0.0130)	(0.0115)		(0.0263)			
YieldWatch×Post		-0.0022	-0.0083	-0.0122	-0.0123		0.0907	0.1729**	0.1561*	0.1052
		(0.0134)	(0.0114)	(0.0104)	(0.0104)		(0.0660)	(0.0764)	(0.0811)	(0.0822)
YieldWatch×Post×Total Offered Yield		-0.0290***	-0.0311***	-0.0313***	-0.0328***		-0.0663**	-0.0934***	-0.0918***	-0.0411
		(0.0095)	(0.0076)	(0.0072)	(0.0066)		(0.0300)	(0.0299)	(0.0276)	(0.0290)
N	94,903	94,029	207,784	303,792	506,960	6,653	6,653	5,035	6,292	7,476
adj. R^2	0.263	0.254	0.390	0.436	0.485	0.268	0.269	0.269	0.272	0.269
Investor controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
T-C Pair ×Week FE	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
T-C Pair ×Farmer FE	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
T-C Pair ×Farm FE	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No
Farm x Week FE	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Week FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Farmer FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Farm FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Weighted Regression	No	No	Yes	Yes	Yes	No	No	No	No	No
# treatment	2,690	2,690	2,690	2,690	2,690	148	148	148	179	197
# control	2,690	2,690	6,982	10,574	17,987	86	86	86	99	128
Sample	nearest 1	nearest 1	nearest 3	nearest 5	nearest 10	[-\$200,\$200]	[-\$200,\$200]	[-\$200,\$200]	[-\$250,\$250]	[-\$300,\$300]
IFO	No	No	No	No	Yes	Yes	Yes	Yes	Yes	
			Unconditional Regression					IFO Threshold Regression		

Table 9: APY.Vision Airdrop Experiment: Information Display and Reaching for Yield

In this table, we investigate whether information on hidden risks and past performance displayed by APY.Vision impacts farmers' reaching for yield propensity. $Flow_{t,t+7}$ defines investors' farm flows over the subsequent 7 days. *APY.Vision NFT token* is an indicator variable equal to one if an investor holds the randomly allocated APY.Vision NFT token in the wallet and zero otherwise. *Total Offered Yield* is the total offered farm yield displayed to investors (APR). The total offered farm yield is the sum of the offered yield and the trading fee yield estimated using the previous day's trading volume. Investor controls include the natural logarithm of the dollar value of the yield farming portfolio of an investor and the natural logarithm of one plus the number of days since the starting date of the yield farming investment. In columns (1) to (3), we compare the 38 wallets that received NFT tokens via airdrops (treated wallets) to all eligible investors on the SushiSwap platform. In columns (4) to (6), we compare treated wallets with all other eligible wallets conditional on satisfying the eligibility requirements described in Appendix Table A.8 and having an investment duration within two weeks of that of a treated wallet. One treated wallet is dropped because it does not have corresponding control wallets. All variables are winsorized at the 0.1% and 99.9% levels. Standard errors are double clustered at the investor and week level. *,**, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A	(1)	(2)	(3)	(4)	(5)	(6)
				$Flow_{t,t+7}$		
Total Offered Yield	0.0285***	0.0186***		0.0277**		
	(0.0048)	(0.0039)		(0.0107)		
APY.Vision NFT		0.0134	0.0101	0.0130	0.0142	0.0141
		(0.0137)	(0.0141)	(0.0162)	(0.0162)	(0.0200)
APY.Vision NFT×Total Offered Yield		-0.0370*	-0.0419**	-0.0504*	-0.0452*	-0.0614**
		(0.0213)	(0.0200)	(0.0276)	(0.0249)	(0.0297)
N	422,700	422,700	422,648	264,317	263,945	263,945
adj. R^2	0.108	0.125	0.137	0.118	0.145	0.559
Include control vars.	Yes	Yes	Yes	Yes	Yes	Yes
T-C Pair ×Week FE	No	No	No	Yes	Yes	Yes
T-C Pair ×Farmer FE	No	No	No	Yes	Yes	Yes
T-C Pair ×Farm FE	No	No	No	Yes	Yes	Yes
Farm × Week FE	No	No	Yes	No	Yes	Yes
Farmer FE	Yes	Yes	Yes	No	No	No
Week FE	Yes	Yes	No	No	No	No
Farm FE	Yes	Yes	No	No	No	No
Weighted Regression	No	No	No	No	No	Yes
# treatment	38	38	38	37	37	37
# control	14,266	14,266	14,266	7,071	7,071	7,071
Sample	All	All	All	All eligible wallets	All eligible wallets	All eligible wallets

INTERNET APPENDIX

Reaching for Yield in Decentralized Financial Markets

Abstract

Yield farms in decentralized finance dynamically compete for liquidity by offering high yields, advertised as salient headline rates. Farming these yields involves complex investment strategies with hidden downside risks. Capitalizing on the transparency of blockchain transactions data, we show that investors chase farms with high yields and that farms with the highest headline rates record the most negative risk-adjusted returns. Through randomized shocks to yield farmers' information display, we show that improved risk disclosure and lower perceived product complexity reduces yield chasing, thereby improving investor performance. Our evidence is consistent with salience theory that may underpin reaching for yield behavior.

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

Keywords: complexity, decentralized finance, derivatives, reaching for yield, salience

A Institutional background

We first provide institutional details on decentralized finance, yield farming, and Binance Smart Chain. We then discuss PancakeSwap and its benefits for studying yield farming.

A.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[9], a public dashboard with data on DeFi, the total dollar value locked (TVL) in decentralized financial services, a measure of market capitalization, increased from less than $1 billion in February 2020 to a peak of $190.82 billion on November 18, 2021, and was measured at $63.55 billion on July 31, 2022 (sample end).

Yield farming is a financial service that offers compensation for liquidity provision in sequential steps. Holders of cryptocurrency tokens can deposit their tokens in liquidity pools, which issue and award 'LP tokens' (a.k.a. 'flip tokens') that certify the liquidity provision and represent a fractional claim on the pool's liquidity. These 'LP tokens' can be deposited in the PancakeSwap main staking contract, which promises farm yields as passive source of income, paid to yield farming investors in the governance token currency of the PancakeSwap platform, called CAKE.

Intuitively, yield farming is a decentralized variant of securities lending, although the chain of transactions is more complex. By offering yield enhancements for liquidity provisions, platform owners (i.e., the aggregate ownership of the native governance tokens) can incentivize liquidity provision. This impacts a platform's long-term success, since, in a decentralized exchange, a more liquid pool implies a smaller price impact per trade, which is desirable for traders. In a lending pool, greater pool liquidity may drive down borrowing costs, which can attract more borrowers. Since the platform owners can vote on the reallocation of yields across farms, they can also channel liquidity to the pools of their choice and encourage adoption of the corresponding tokens.

Headline rates for promised investment performance 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 5.

While yield farming is marketed as being simple through means of engaging platforms, cartoons, rockets, and emojis, both executing a yield farming investment and understanding its payoffs is complex. The strategy involves a sequence of 12 transactions in three different underlyings. Return performance is highly non-linear and comes from 4 components: realized farm yield from staking LP tokens, capital gains from cryptocurrencies staked to liquidity pools, fees from trading by third-party investors in liquidity pools, and impermanent losses driven by relative price changes of the cryptocurrencies locked in liquidity pools. Thus, the complexity of yield farming resembles obfuscated investment strategies observed 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 during our sample period. Transaction costs in PancakeSwap are significantly lower than in other popular decentralized exchanges like Uniswap (Figure 2). This lowers the barriers to entry for retail investors, who are active investors in yield farms.

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

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 taken by the U.S. Securities and Exchange Commission, who have become increasingly vocal about enhanced regulatory scrutiny of decentralized financial services (e.g., Gensler, 2021). Our work is intended to inform this ongoing debate by means of assessing the risk and return characteristics of yield farming strategies.

A.2 Binance Smart Chain

Binance Chain was launched by Binance in April 2019.[10] Its main goal is to facilitate 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 significant 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 for the dual purpose of maintaining the high throughput of Binance Chain and allowing the integration of smart contracts.

In the BSC ecosystem, Binance Coin (BNB) is used as the basic medium of exchange, similar to Ether (ETH) in the Ethereum network. End users pay their transaction fees in BNB and use BNB to trade cryptocurrencies on 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, BSC allows for a maximum of 300 transactions per second.[11] 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 lower than those of Ethereum. As of September 5, 2021, the average transaction fee charged by BSC is \$0.399, whereas it is \$5.842 for Ethereum. 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.[12]

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.[13] As of the same date, Binance Coin is 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

[10] In February 2022, Binance Smart Chain rebranded to BNB Smart Chain.

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

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

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

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.

A.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 (2024).

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 0.20% and 0.25% of the trading volume in versions v1 and v2, respectively, of which 0.17% is added to the liquidity pool associated with the corresponding cryptocurrency pair. Despite the earnings potential, investors are also exposed to price risk associated with impermanent losses, which are driven by return divergence across a pool's tokens.

In addition to the income generated from trading fees, liquidity providers can passively earn 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 in PancakeSwap's main staking contract. Farm yields are paid in PancakeSwap's governance token, called CAKE.

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. In April 2023, PancakeSwap announced another migration to version 3, which was concluded by April 20. We study yield farming for versions v1 and v2.

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

A.4 PancakeSwap as an ideal laboratory to study yield farming

Many decentralized trading venues offer passive income opportunities through yield farming. Among DeFi platforms, Uniswap and PancakeSwap consistently lead the league ranks in terms of trading activity in our sample period. The key difference between both platforms is that Uniswap (PancakeSwap) runs on the Ethereum blockchain (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 reward. However, there are no farms in Uniswap to which liquidity providers can stake their LP tokens to earn additional income through yield farming.

Second, PancakeSwap is one of the largest decentralized exchanges. In Table A.3, we report the daily trading volume for the ten largest decentralized exchanges as of October 9, 2021.[14] The largest DEX is dYdX, which specializes in derivatives trading. Augustin, Rubtsov, and Shin (2023) 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 A.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, PancakeSwap registered 435,130 active users on October 24, 2021, in contrast to 47,730 active users recorded for Uniswap.[15] 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 with heterogeneity in yield farming opportunities. This provides important variation to help understand the risk and return characteristics of yield farms. We study 262 unique yield farms that were active between the inception of PancakeSwap on September 23, 2020 and July 31, 2022.

[14] The data in Table A.3 are from Coinmarketcap. Also according to DeFiLlama, PancakeSwap is ranked second next to Uniswap in daily trading volume and cumulative trading volume as of March 15, 2024.

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

B Derivation of conceptual framework

In this section, we provide supporting explanations for the conceptual framework and detailed steps in the derivation of Equations (8) and (10). In Section B.1, we ignore frictions, which we cover in Section B.2.

B.1 Capital gains and impermanent loss

To help build intuition, we explain all derivations using a specific example. We assume the existence of a liquidity pool with a cryptocurrency pair $(A, B) \equiv (BNB, BUSD)$. Thus, this liquidity pool covers the BNB-BUSD cryptocurrency token pair, where BUSD is a stablecoin pegged to USD. Assuming a BNB–BUSD exchange rate of 100, a liquidity provider deposits 1 BNB and 100 BUSD at time t to the liquidity pool. After the liquidity provision, the aggregate liquidity in the pool is 10 BNB and 1,000 BUSD, implying that the liquidity provider's fractional owenrship is 10%. After h days, at time $t+h$, the BNB price increases to, for example, 200 BUSD (hypothetically). The liquidity provider withdraws his/her liquidity.

The constant product model imposes that the product of the aggregate number of tokens in the pool is equal to constant K, i.e., $k = \alpha_t^A \cdot \alpha_t^B = 10 \times 1,000 = 10,000$, where α^i denotes the number of tokens of cryptocurrency i in the liquidity pool. Lemma B.1 shows that the valuation of token A (i.e., BNB) should be identical to the valuation of token B (i.e., BUSD) at any t, i.e., $\alpha_t^A \cdot P_t^A = \alpha_t^B \cdot P_t^B$ for all t.

Lemma 1. *In a constant product automated market maker, $\alpha_t^A \cdot P_t^A = \alpha_t^B \cdot P_t^B$ for all t.*

Proof. Under the constant product model, the product of the quantities of two cryptocurrencies should be constant, i.e. $\alpha_t^A \cdot \alpha_t^B = k$. This implies that $\frac{\partial \alpha_t^B}{\partial \alpha_t^A} = -\frac{\alpha_t^B}{\alpha_t^A}$. A third-party investor wanting to purchase δ units of A for the sale of asset B would need to sell a quantity B equivalent to $\delta \frac{\alpha_t^B}{\alpha_t^A}$. This implies that $\delta \cdot P_t^A = \delta \frac{\alpha_t^B}{\alpha_t^A} \cdot P_t^B \rightarrow P_t^A \alpha_t^A = P_t^B \alpha_t^B$. $\square$

Since we have two equations including the aggregate number of tokens A and B, $\alpha_t^A \cdot P_t^A = \alpha_t^B \cdot P_t^B$ and $k = \alpha_t^A \cdot \alpha_t^B$, we can solve for the expressions of α_t^A and α_t^B, such that:

$$\alpha_t^A = \sqrt{k \left(\frac{P_t^B}{P_t^A} \right)}, \quad \alpha_t^B = \sqrt{k \left(\frac{P_t^A}{P_t^B} \right)}. \tag{B.1}$$

We numerically illustrate the impact of a transaction by a third-party investor on the pool's token composition at time $t+h$ using the example of an increase in the exchange rate of BNB–BUSD from 100 to 200, which is equivalent to \$100 to \$200 if we assume that BUSD is perfectly pegged to USD:

$$\alpha_{t+h}^A = \sqrt{k \left(\frac{P_{t+h}^B}{P_{t+h}^A} \right)} = \sqrt{10,000 \times (\$1/\$200)} = \sqrt{50} = 7.07,$$
$$\alpha_{t+h}^B = \sqrt{k \left(\frac{P_{t+h}^A}{P_{t+h}^B} \right)} = \sqrt{10,000 \times (\$200/\$1)} = \sqrt{2,000,000} = 1414.21.$$

The liquidity provider's fractional pool ownership is 10%. Upon redemption, he/she will receive 10% of the pool's tokens, corresponding to 0.707 BNB and 141.421 BUSD. This amounts to $0.707 \times 200 + 141.421 \times 1 = \$282.82$.

Compare the redemption value to the counterfactual buy-and-hold strategy of the two tokens (1 BNB and 100 BUSD). In that case, the liquidity provider's portfolio would be worth $\$300 = 1 \times 200 + 100 \times 1$, more than the redemption value after liquidity provision. The difference is the impermanent loss, which arises due to divergence in price correlation of tokens A and B. In this case, the impermanent loss corresponds to a loss of $(282.82/300 - 1) \times 100 =$ -5.727%.

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 corresponds numerically to $(282.82/300 - 1) \times 100 = -5.727\%$.

We formalize the impermanent loss through the ratio of the portfolio value in the liquidity provision and buy-and-hold strategies minus one, using a generic ownership share ω:

$$\frac{\omega(P^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h})}{\omega(P^A_{t+h}\alpha^A_t + P^B_{t+h}\alpha^B_t)} - 1. \tag{B.2}$$

We emphasize that α^i in the denominator corresponds to the number of tokens in the intial liquidity provision, whereas α^i in the numerator corresponds to the number of tokens after trading by third-party investors between t and $t+h$ has changed the token composition in the pool. We rewrite Equation (B.2) in terms of price ratios P^A/P^B and P^B/P^A using Equation (B.1):

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

We can simplify the above expression using the relative price ratio $\rho_t = \frac{P^A_t}{P^B_t}$:

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

The above expression illustrates the impermanent loss as a function of the relative price ratio between two tokens. This clearly emphasizes that, as long as prices are perfectly correlated, i.e., $\rho = 1$, there will be no impermanent loss. As soon as $\rho \neq 1$, there is a loss, since it is straightforward to show that the impermanent loss is strictly non-positive, i.e., $\frac{2\sqrt{\rho_{t+h}/\rho_t}}{\rho_{t+h}/\rho_t+1} - 1 = -\frac{(\sqrt{\rho_{t+h}/\rho_t}-1)^2}{\rho_{t+h}/\rho_t+1} < 0$. Figure A.5 illustrates numerically the non-linearity between the impermanent loss and ρ_{t+h}/ρ_t.

For our analysis, we simplify the liquidity provider's gross return defined as:

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

by decomposing it into two independent parts:

$$\begin{aligned}
\frac{\omega(P^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h})}{\omega(P^A_t\alpha^A_t + P^B_t\alpha^B_t)} &= \underbrace{\left(\frac{P^A_t\alpha^A_t}{P^A_t\alpha^A_t + P^B_t\alpha^B_t}\right)R^A_{t,t+h} + \left(\frac{P^B_t\alpha^B_t}{P^B_t\alpha^A_t + P^B_t\alpha^B_t}\right)R^B_{t,t+h}}_{Capital\ Gains} \\
&+ \underbrace{\frac{P^A_{t+h}\alpha^A_{t+h} + P^B_{t+h}\alpha^B_{t+h}}{P^A_t\alpha^A_t + P^B_t\alpha^B_t} - \left[\left(\frac{P^A_t\alpha^A_t}{P^A_t\alpha^A_t + P^B_t\alpha^B_t}\right)R^A_{t,t+h} + \left(\frac{P^B_t\alpha^B_t}{P^B_t\alpha^A_t + P^B_t\alpha^B_t}\right)R^B_{t,t+h}\right]}_{Impermanent\ Loss},
\end{aligned}$$

where the gross returns to tokens A and B are defined as $R^A_{t,t+h} = \left(P^A_{t+h}\alpha^A_{t+h}\right) / \left(P^A_t\alpha^A_t\right)$ and $R^B_{t,t+h} = \left(P^B_{t+h}\alpha^B_{t+h}\right) / \left(P^B_t\alpha^B_t\right)$. The first term, which we call capital gains, reflects the counterfactual return from a buy-and-hold investment strategy without liquidity provision to the pool. The second term defines the impermanent loss and reflects the return difference between the liquidity provision and a buy-and-hold strategy.

Using Lemma B.1, we can rewrite the expression for capital gains as:

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

Using Lemma B.1, we can also simplify the expression for the impermanent loss as:

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

It is straightforward to show that the impermanent loss defined in the context of return on liquidity provision is closely related to the percentage impermanent loss:

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

B.2 Trading frictions in yield farming

We examine three types of trading frictions: gas fees, trading fees, and price impact.

Gas fees: Table A.2 lists the 14 steps involved in one round trip yield farming strategy. Among the 14 steps, 10 require the farmers to pay gas fees. Gas fees are the transaction costs imposed on BSC users for using the computational resources of the network. Gas fees are flat overhead costs and are not proportional to the size of the transaction. We source gas fees from Bitquery, a data provider specialized in blockchain services. To compute the returns to yield farming with frictions, we subtract the aggregate gas fee for each round trip investment from the initially invested capital.

Trading fees: Let c^*=0.0025 (0.25%) denote the trading fee cost paid by third-party investors as a proportion of trading volume. In step 2 of Table A.2, the purchase of token A requires the payment of a 0.25% trading fee. Because this trading fee applies to half the investment amount dedicated to token A, the farmer effectively pays half the trading fee $\frac{c*}{2}$ (=0.125%). Moreover, the farmer pays an additional $\frac{c*}{2}$ fee when he/she converts the withdrawn token A to token B. Similar arguments apply in steps 3 and 13 enumerated in Table A.2. A yield farmer would thus pay $\frac{c*}{2}$ four times, implying that a farmer's gross return on capital gain and impermanent loss should be adjusted by $(1 - 2c^*)$.

In step 10 of Table A.2, the yield farmer also pays trading fees when he/she sells CAKE tokens harvested from yield farming. Thus, we further multiply the realized farm yield term in Equation (8) with $(1 - c^*)$.

Price impact: Executing a yield farming strategy involves buying and selling token A, as we illustrate in step 2 of Table A.2. 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 losses for the yield farmer. The size of the loss is proportional to the relative contribution of the investment (I_t) to the size of the liquidity pool, i.e., $I_t = f \cdot L_t$. We go through each step in Table A.2 to examine the price impacts involved in a yield farming strategy.

(1) Step 1: The liquidity pool has two tokens A and B. The aggregate number of tokens are given by α_t^A and α_t^B and their prices are denoted by P_t^A and P_t^B.

(2) Step 2: A yield farmer must provide tokens A and B in equal amounts. Thus, he/she must acquire tokens A and B proportionally to α_t^A/α_t^B. For this purpose, we divide his/her investment into $x \cdot I_t$ and $(1-x) \cdot I_t$ to allocate towards tokens A and B, respectively. The yield farmer first converts $\$x \cdot I_t$ to acquire token B in a liquid market for B. Then, the farmer will own $x \cdot \frac{I_t}{P_t^B}$ of token B, which he/she will use to buy Δ_t^A units of token A by means of the liquidity pool. Due to the constant product model technology, we have that:

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

Solving for Δ_t^A yields:

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

(3) Step 3: The yield farmer uses the remaining 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 solve for x that satisfies: $\frac{\Delta_t^A}{\Delta_t^B} = \frac{\alpha_t^A}{\alpha_t^B}$.

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

implying that:

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

with two solutions for x, with the relevant positive solution given by:

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

(4) Step 4: Arbitrageurs correct the price by supplying Δ_t^A of token A in reurn for Δ_t^B of token B. This restores the liquidity pool to its initial state.

(5) Step 5: The yield farmer receives LP tokens to certify the liquidity provision. 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:

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

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

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

(7) Step 7: The yield farmer waits for h days. After trading by third-party investor, the aggregate number of tokens A and B in the pool change and become $\alpha_{t+h}^A(1+s(f))$ and $\alpha_{t+h}^B(1+s(f))$.

(8) Step 8: The yield farmer receives (harvests) 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 has withdrawn liquidity, the shares of token A and B in the pool change to α_{t+h}^A and α_{t+h}^B.

(12) Step 12: The yield farmer sells his/her $\Delta_{t+h}^A = s(f)\alpha_{t+h}^A$ of token A and receives Δ_{t+h}^B of token B. At this stage, there are α_{t+h}^A and $\alpha_{t+h}^B$of token A and token B in the pool. After the farmer has sent $\Delta_{t+h}^A = s(f)\alpha_{t+h}^A$ of token A, he/she receives Δ_{t+h}^B units of token B. Due to the constant product model, we have that:

$$\left(\alpha_{t+h}^A + s(f)\alpha_{t+h}^A\right)\left(\alpha_{t+h}^B - \Delta_{t+h}^B\right) = \alpha_{t+h}^A\alpha_{t+h}^B \rightarrow \Delta_{t+h}^B = \frac{s(f)}{1+s(f)}\alpha_{t+h}^B.$$

The farmer sends $s(f)\alpha_{t+h}^A$ units of token A in return for $P_{t+h}^B\Delta_{t+h}^B$ worth of USD. Thus, the effective price faced by the yield farmer when selling token A is equal to:

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

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

(13) Step 13: The yield farmer sells $\Delta_{t+h}^B + s(f)\alpha_{t+h}^B$ units 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 in reurn for Δ_{t+h}^A units of token A. A new round of yield farming starts.

Our goal is to 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 $s(f)\alpha_t^B$ units of token A and B at $\tilde{P}_t^A$ and P_t^B. After h days, the yield farmer withdraws $s(f)\alpha_{t+h}^A$ and $s(f)\alpha_{t+h}^B$ units of token A and B and sells them at $\tilde{P}_{t+h}^A$ and P_{t+h}^B. In this case, the gross return can be expressed as:

$$\frac{\tilde{P}_{t+h}^A\left(s(f)\alpha_{t+h}^A\right) + P_{t+h}^B(s(f)\alpha_{t+h}^B)}{\tilde{P}_t^A\left(s(f)\alpha_t^A\right) + P_t^B(s(f)\alpha_t^B)} = \frac{\tilde{P}_{t+h}^A\alpha_{t+h}^A + P_{t+h}^B\alpha_{t+h}^B}{\tilde{P}_t^A\alpha_t^A + P_t^B\alpha_t^B}.$$

We simplify this expression as follows:

$$\frac{\tilde{P}^A_{t+h}\alpha^A_{t+h}+P^B_{t+h}\alpha^B_{t+h}}{\tilde{P}^A_t\alpha^A_t+P^B_t\alpha^B_t}=\frac{\left(\frac{1}{1+s(f)}\right)P^A_{t+h}\alpha^A_{t+h}+P^B_{t+h}\alpha^B_{t+h}}{P^A_t\left(1+\left(f-1+\sqrt{f^2+1}\right)\right)\alpha^A_t+P^B_t\alpha^B_t}$$

$$=\frac{\left(\frac{1}{1+s(f)}+1\right)P^A_{t+h}\alpha^A_{t+h}}{\left(1+\left(f-1+\sqrt{f^2+1}\right)+1\right)P^A_t\alpha^A_t}=\frac{\frac{1}{1+s(f)}+1}{f+1+\sqrt{f^2+1}}\left(\frac{P^A_{t+h}\alpha^A_{t+h}}{P^A_t\alpha^A_t}\right)$$

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

where

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

Accounting for both price impact and trading fees, the gross return is adjusted as follows:

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

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

Figure A.6 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.

Aggregate Frictions for a Given Holding Time:

We assume that investors arrive uniformly and hold their positions for a period of τ each time. Therefore, frictions are incurred h/τ times in aggregate over a horizon h. We approximate the aggregate loss from frictions as: $\left(R_{t,t+h}-R^{friction}_{t,t+h}\right)\times\frac{h}{\tau}$.

C Data appendix

We describe the technical construction of all data sets related to farms (C.1), prices (C.2), token transfers (C.3), Yieldwatch (C.4), APY.Vision Giveaways (C.5), and cryptocurrency factors (C.6). We also discuss the cleaning of farm (C.7) and farmer (C.8) data.

We source all information from public blockchains. Interactions with the blockchain are facilitated by the Web3 application programming interface (API) together with a blockchain archive node service. Archive nodes provide the record of all blocks since inception of the blockchain, unlike full nodes, which tend to store only more recent blockchain data. The Web3 API operates as a middleman between the user and the archive node, allowing for a simple object-oriented programming interface. Our technical documentation primarily reflects the perspective of a

Python Web3 interface, although most of its functionality should be comparable to Web3 interfaces based on other programming languages.

To clarify the time convention, note that blockchain time is measured by the 'height' of a block, a unique identifier which represents the order of a particular block relative to the inception of the blockchain. The history of all mined blocks is freely accessible through an archive node or a block explorer service. A given block height can be passed onto an archive node as an optional argument in an API to 'restrict' observable data to everything in that block and before. This enables the reproduction of the historical conditions of the blockchain corresponding to a specific point in time. For our analysis, we use the height of the last block mined on each day in coordinated universal time (UTC+0) to index an observation on a particular date for all following sections. As an example, the last Binance Smart Chain block height in our sample is 20045095, a block that was mined on July 31, 2022 at 11:59:58 PM UTC, and used to extract observations for July 31, 2022.

C.1 Accessing farm data

We extract farm data in two steps: first, we identify each day the set of addressable smart contracts corresponding to yield farms on a given platform and calculate, for each farm, the quantities of interest in the set.

We find the set of contracts on each day by interacting with the active version of the main staking contract through the $poolInfo(q)$ function, where q is a non-negative integer. This returns information about the q-th liquidity pool added to the main staking contract: first, the blockchain address of the pool, and second, the weight corresponding to the current share of minted tokens that this pool receives ('allocPoint'). The total amount of pools stored in this way is given by the $poolLength()$ function at each day, and it is straightforward to iterate over them until all information has been collected.

For each farm, we then make several direct calls to its smart contract to extract further information: $token0()$ and $token1()$ return the addresses of the two tokens traded by the pool, while $getReserves()$ returns the balances of the two tokens. To stake tokens, users must transfer LP tokens to the main staking contract, so we simply call $balanceOf(mainStakingContract)$ to get the amount of staked tokens, and $totalSupply()$ to get the total outstanding amount of LP tokens. For these steps, it is critical to account for variable decimal precision across different tokens, which we can identify by calling the $decimals()$ function for each token and then explicitly adjust for.

C.2 Prices and trades

To identify prices for each token, we use the $getAmountsOut(N, [B, Token])$ function from the main router of the platform. The main router for a decentralized exchange is the smart contract responsible for quoting swap rates, and $getAmountsOut(N, [B, Token])$ is a request for a quote where N units of token B are exchanged for quoted units of $Token$, absent considerations of fees or price impact. For our purpose, we take the wrapped version of the native token for a block chain (e.g. WBNB for BSC) as B, and 0.01 as N.

Our choice of B is motivated by the fact that the native token is the most liquid token on its smart chain, N is chosen such that there is no loss of exchange rate precision from trading overly small amounts[16], nor is the size big enough to distort the liquidity of any pool. This function allows us to compute exchange rates on the decentralized exchange. Finally, we determine the exchange rate between B and $USDT$ and source a centralized quote for $USDT$ against USD. This allows us to get the exchange rates of all tokens versus USD.

[16] If N is less than 1e-06 USDT, the quote is zero because it is less than the token's decimal precision.

To determine trading volume for each pool, we pass the pool address and date range to Bitquery, a third-party data provider which calculates the total amounts of the two tokens traded through a given smart contract for a given date range. We merge this information with the main dataset. Once all of these items are compiled, it is then straightforward to calculate the total return and its individual components according to our formulae in the main text: the offered yield, the offered total yield, and the staking ratio.

C.3 Token transfer data

We construct token transfer data using the event logs emitted by smart contracts when they update their internally stored variables. We utilize the *eth.getLogs*() function to collect these logs. Each log has the following fields: the block height of record, the smart contract/token which performs the update, an id corresponding to the particular event that has occurred (the 1st 'topic'), additional key information (subsequent 'topics'), and additional general information (log 'data').

Events for which LP tokens are transferred always follow the same event id, which allows us to restrict our attention to the set of events corresponding to this particular id and emitted by the LP tokens within our sample. For such events, the second topic in an event log is the address of the token sender, and the third topic is the recipient. The amount transferred is contained in the log data, in hexadecimal.

We validate a subset of the reconstructed token transfer data against the displayed token transfer records on the bscscan website (https://bscscan.com/), and find a perfect match. These data form the backbone of our user-level analysis.

C.4 Yieldwatch Initial Farm Offering

Yieldwatch is a smart dashboard hosted on the Binance Smart Chain. Through its user interface, Yieldwatch provides enhanced information on yield farming and token staking performance, including information related to pool liquidity, impermanent loss, pending yield rewards and generated trading fees.

Access to Yieldwatch Pro, Yieldwatch.net's main service, is granted through the acquisition of Yieldwatch tokens (henceforth WATCH tokens). Yieldwatch organized an initial farm offering (IFO) of WATCH tokens at PancakeSwap on March 4, 2021.[17] The IFO was organized during one hour, between 5p.m. and 6p.m. SGT on 4th of March 2021, corresponding to the sale start and end blocks 5383360 and 5384550, respectively.

A total of 8 million WATCH tokens (40% of maximum supply) was supplied during the IFO at a fixed offer price of $0.1 per WATCH token. Thus, the intention was to raise $800,000 USD worth of WATCH tokens, for which investors could bid using CAKE-BNB LP tokens. Final allocation of WATCH tokens was administered using the so-called "overflow" sales method. According to PancakeSwap's Medium page, this implies that providing more funds would lead to a greater token acquisition and that, in the event of oversubscription, individual allocations would be be prorated according to the percentage of the aggregate funding bids.[18]

As an example, consider a total supply of $8 million for a fixed supply of $800,000 worth of WATCH tokens. This corresponds to an oversubscription of 10 times the target fundraising amount. An individual bid of $1,000 corresponds

[17] Since March 4 is immediately after the start of our sample period, we also use data before March 1, 2021 for the reaching for yield analysis.

[18] See `https://medium.com/pancakeswap/yieldwatch-watch-ifo-to-be-hosted-on-pancakeswap-d24301f17241`.

to 0.0125% of the aggregate capital supply, which, on a pro-rated basis, would give the bidder a right to purchase 1,000 WATCH tokens at a price of $100 (i.e., $0.1 per token). In that case, 90% of the supplied capital or $900 would be returned to the bidder.

The IFO attracted 16,061 users bidding a total of $569 million for the 8 million WATCH tokens. The first subscription time was at 05:01:16 p.m. SGT and the last subscription time was at 6:01:12 p.m. SGT. Importantly, both the aggregate capital supply and the initial threshold for token acquisition were unpredictable. This gives rise to a threshold analysis with quasi-randomized token allocation.

To support the quasi-random nature of token allocation around a threshold, we provide in Panel A of Figure A.11 the ex-post allocation schedule of WATCH tokens. In contrast to the announced continuous proportionality rule, this figure shows that the allocation of tokens was implemented as a step function by multiples of 8 tokens. Moreover, even if someone would have been aware of the details of the allocation mechanism, it would have been impossible to perfectly predict the aggregate capital supply in the IFO. In fact, Figure A.12 shows that the dynamics of the capital supply in the Yieldwatch IFO do not exhibit strategic bidding patterns. We further verified that those investors who received token allocations did not strategically delay their bids to the final minutes of the IFO.

Panel B of Figure A.11 focuses on the first allocation cut-off by restricting the x-axis to capital supply below $1,200. Investors who supplied less than $569.4 did not receive any WATCH tokens while investors who supplied more than $569.4 received 8 WATCH tokens. As discussed, it is virtually impossible to predict the aggregate capital supply and the details of the allocation rule were not transparent. The lack of evidence on strategic bidding further mitigates concerns that the acquisition of WATCH tokens around the first threshold is correlated with investor skill. As a result, we conclude that the allocation of WATCH tokens around this threshold is quasi-random and that investors bidding just below and just above this threshold are similar.

C.5 APY.Vision Giveaways

In our sample period, APY.Vision organized several airdrops. Airdrops (or giveaways) are events in which APY.Vision gives a select group of users access to premium tracking services. APY.Vision operates across multiple platforms and selects users randomly to provide their services. We identify 20 airdrops through their announcements on X (formerly Twitter) and the bulk-transfer patterns of NFT tokens on specific days between December 2020 and May 2022.

Upon selection, a recipient is granted a unique NFT belonging to a collection of NFTs. Each collection is unique to a specific giveaway. Using NFT tracking websites Rarible and Opensea, we first match each giveaway to the corresponding collections minted by the APY Vision minting address and then collect information on the address of the recipient manually.

Each giveaway has unique eligibility requirements for receiving NFTs. The announced number of wallets is selected randomly from those that satisfy the requirements. Detailed information about the giveaways is presented in Table A.8. Out of 20 giveaways, 16 of them present eligibility requirements verifiable from public blockchain data, and four of them do not. These four are "CoinGecko + APY Vision #1", "CoinGecko + APY Vision #2", "APY Vision V2 Release Party NFT", and "APY Vision Ambassador NFT" airdrops.

For the 16 giveaways, we generally follow the described requirements strictly. Ideally, we should find that all the NFT recipients satisfy the eligibility requirements described on the giveaway website. However, the information on the website is sometimes ambiguous, and some NFT recipients do not strictly satisfy the requirements. For those cases, it is unavoidable to make assumptions regarding the requirements for eligibility. For example, the eligibility requirement for "Balancer + APY Vision" is that investors should hold veBAL tokens on April 7, 2022, 10 PM UTC.

The number of NFT recipients from this airdrop was 30. However, out of 30 recipients, 23 recipients strictly meet the requirement. If we relax the requirement by checking the balance on April 8, 2022, at 9 PM instead, because the NFT started to be airdropped from 09:36:11 PM, we find that all 30 recipients satisfy the requirement. We transparently describe detailed information about the necessary adjustment under the column header "Procedure for collecting data" in Table A.8. Following this procedure for data collection, we find most of the NFT recipients satisfy the requirements: Out of 321 NFT recipients from the 16 giveaways, 306 of them, which is 95% of 321, satisfy the requirements. Among them, 35 are in our Sushiswap sample data.

From the four giveaways where the eligibility requirements are not verifiable from the blockchain data, we identified 291 wallets that received NFTs. Due to the lack of data on eligible wallets, it is unavoidable to make an assumption about eligible wallets. Therefore, we assume that all wallets are eligible for receiving the NFTs from the four giveaways. Among the 291 wallets, 3 are in our SushiSwap sample data. Overall, we identified 38 NFT recipients in total, on which our analysis is based.

C.6 Cryptocurrency Factors

Liu, Tsyvinski, and Wu (2022) 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. We then compute volume-weighted average prices across all markets for which Coinmarketcap has data. Our risk-free rate is from the St. Louis Fed's one-month constant maturity Treasury rate.

We exclude from our sample coins without trading volume, 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. All returns are measured in U.S. dollars. The daily excess cryptocurrency market return is constructed as a value-weighted portfolio of all coins with data on the portfolio formation day (prior to applying the filters) minus the risk-free rate.

The excess cryptocurrency size factor is computed using the return from a long-short trading strategy that takes a long (short) position in the value-weighted portfolio of coins ranked in the bottom (top) quintile of market capitalizations on the portfolio formation day. For the cryptocurrency momentum factor, we exclude coins for which the three-week price history is unavailable. The momentum factor is then constructed from a long-short strategy with a long (short) position in the value-weighted portfolio of coins ranked in the top (bottom) quintile of coins with positive three-week momentum on the portfolio formation day.

As a test of the accuracy of our methodology, we replicate the three-factor regressions from Table 11 in Liu, Tsyvinski, and Wu (2022) for portfolios sorted on one-week momentum by quintile, a set of implementable trading strategies not used in the construction of the three factors. Table A.6 provides summary statistics on the coins used for the construction of cryptocurrency factors. In Table A.5, we compare our parameter estimates to those obtained in Liu, Tsyvinski, and Wu (2022) (NBER version dated May 2019, strategy "r 1,0" in Table 11). The two are nearly identical with only minor deviations, which may be due to small variations in the sample period used and/or changes in the markets for which Coinmarketcap tracks price data.[19]

[19] See Schwenkler, Shah, and Yang (2023) for recent work on the reliability of cryptocurrency data.

In addition, it is worth noting that the estimates for alpha obtained in Liu, Tsyvinski, and Wu (2022) 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.5, 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 (2022), in which three-factor weekly alphas exceed 0.02 (or an annualized alpha of 1.80) for many price- and momentum-based strategies.

C.7 Data Cleaning: Farm Data

After the initial data construction, we identify 304 unique yield farms. We restrict our analysis to the sample period March 1, 2021 to July 31, 2022. This reduces our sample to 299 farms with 281 unique cryptocurrencies and 61,023 observations. We exclude 6 farms where UST was one of the tokens in the LP pool (UST-MCOIN, 'UST-MIR', 'UST-mAMZN', 'UST-mGOOGL', 'UST-mNFLX', and 'UST-mTSLA'), since we could not obtain a reliable estimate of the token price for the complete sample, owing to the events of the Terra-Luna collapse. We further exclude two farms, 'PNT-PBTC' and 'QSD-KUN', for which our data provider does not have any trading volume information. We exclude farms that have less than two weeks' worth of data in our sample. These farms most likely correspond to 'Farm Auctions' - promotional partnerships where a liquidity pool receives yield for a week to generate interest and trading activity. These filters reduce our sample to 262 unique farms that cover 247 unique cryptocurrencies and 59,051 daily observations at the farm level.

C.8 Data Cleaning: Farmer Data

We collect all wallet addresses with transactions in the 262 unique farms in our sample. After removing two common burn (null) addresses that do not represent investors, we have 1,190,623 unique wallet addresses (wallets) which have provided liquidity to 529 unique smart contracts (i.e., liquidity pools), corresponding to 2,687,061 unique wallet-liquidity pool pairings (wallet pools) and 62,352,957 observed historical states of wallet pools (total positions).

In a second step, we exclude wallets which trade at an implausibly high frequency, with over 10,000 positions held through our sample period. This reduces our sample to 1,190,442 wallets in 529 liquidity pools and 2,682,603 wallet pools, and 23,944,735 total positions.

Third, we exclude wallets with a smart contract interface, since their positions are not directly managed by investors. This lowers our sample to 1,172,762 wallets in 529 liquidity pools and 2,647,640 wallet pools, and 18,618,936 total positions.

Finally, we exclude wallet-pools in which the wallet transferred LP tokens to a third party, as these represent multi-platform strategies outside the scope of this paper. This lowers our sample to 641,477 wallets in 529 liquidity pools and 1,442,486 wallet pools, and 7,838,261 total positions.

To implement our analysis at the investor level, we need to collect additional information that leads to further loss in observations. We create two separate data sets to implement analysis on flows and on returns.

For investor-level flow analysis, we define flow based on the balance of # LP tokens measured on every Sunday midnight at the weekly frequency. Therefore, wallets that do not have positions as of Sunday midnight are dropped from our sample. This selection process yields 446,227 wallets corresponding to 5,975,858 positions for investor-level flow analysis. These wallets make 10,818,661 transactions in our raw dataset.

For our investor-level return analysis, we make a few additional refinements to reasonably estimate the investor returns. First, we discard positions worth less than \$1 at the start of the holding period, since investors must pay gas fees in excess of 100% of their initial investment to liquidate these positions. Secondly, we require that all positions held by the yield farmer are entered and exited during the yield farm's active period, to ensure that each investor-level return is an unbiased representation of the true yield farming return realized by the respective investor. After these additional refinements, we are left with 532,713 investor return observations. Finally, we focus on investors whose total participation in yield farming lasts at least one day, since our farm-level data is at daily frequency. This leaves 439,639 investors corresponding to 6,183,222 positions in our final investor-level return analysis. These wallets make up for 11,175,192 transactions in our raw dataset.

Figure A.1: 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 natural logarithm of the dollar value of liquidity in the yield farm in units of \$1 million. The y-axis corresponds to the natural 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 natural logarithm of (1 + annualized offered farm yield) on the natural logarithm of the dollar value of liquidity in the yield farm.

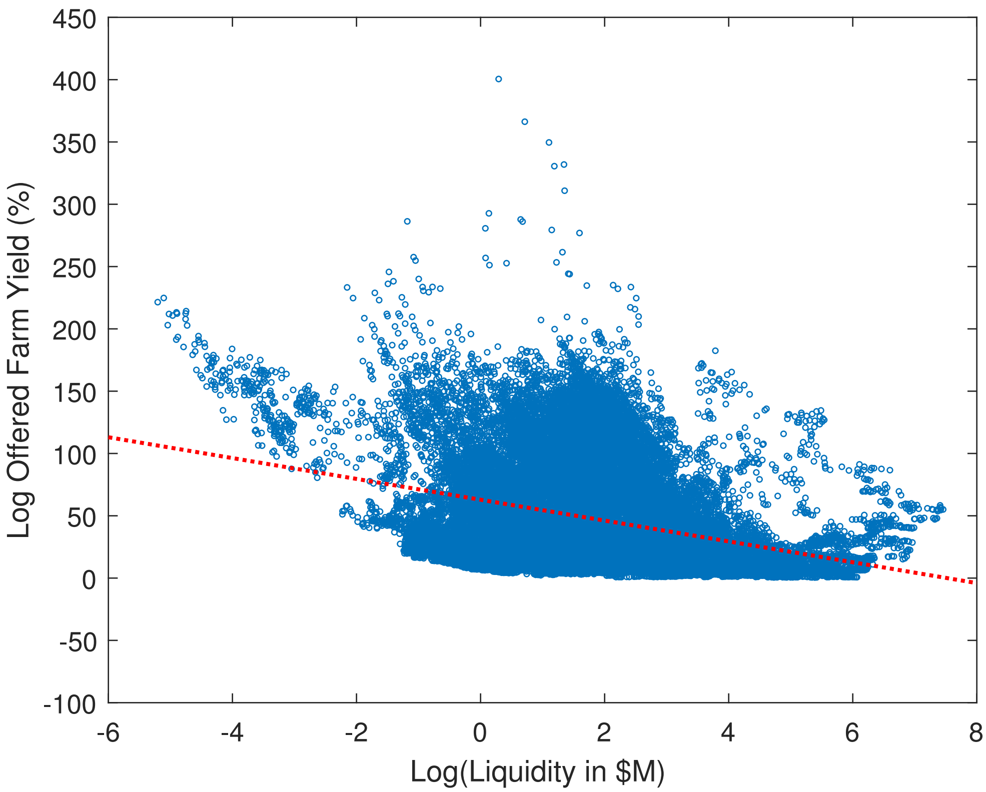


Figure A.2: Relation between Model-implied and Listed Offered Farm Yields

In this figure, we compare the offered farm yields calculated using Equation (6) on the y-axis to those listed on PancakeSwap's homepage on the x-axis (https://PancakeSwap.finance/farms). The listed farm yields were 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 $\widehat{y}_t = 1.002 \times y_t - 0.001$.

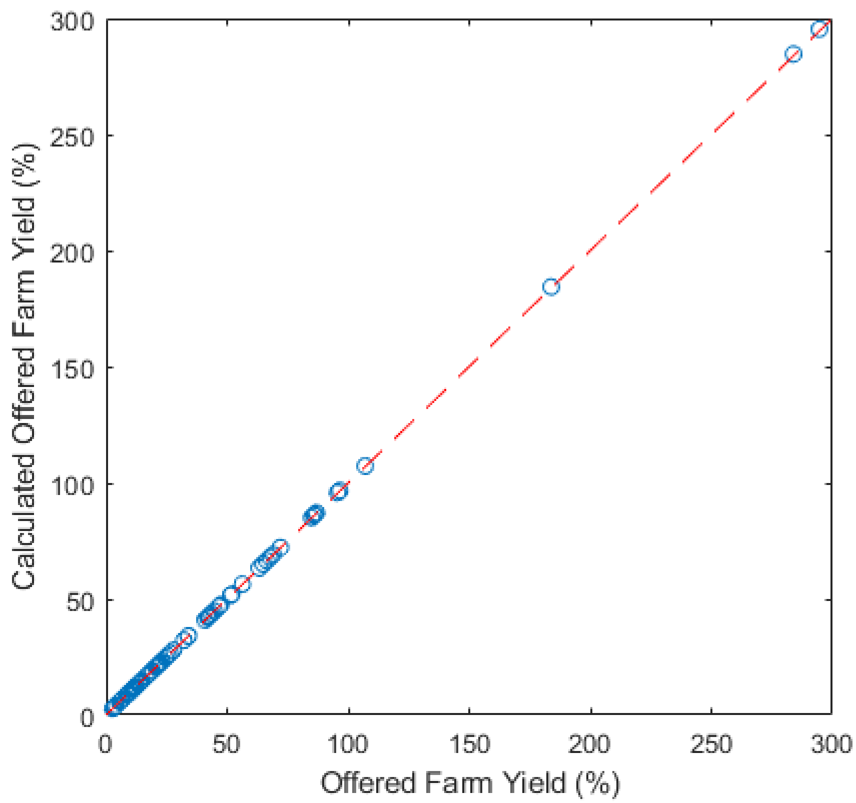


Figure A.3: User Interface of Yield Farms in PancakeSwap

In this figure, we provide a snapshot of the user-interface environment for yield farms in PancakeSwap. For a current snapshot, see https://PancakeSwap.finance/farms.

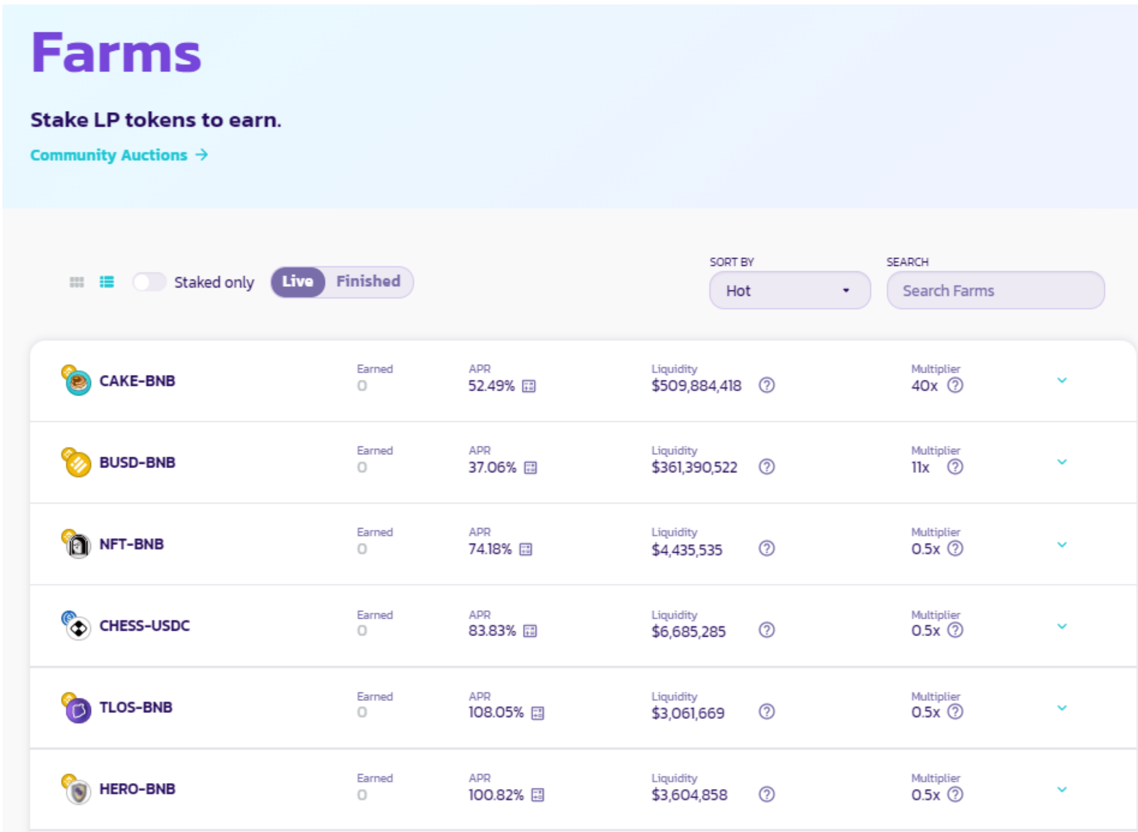


Figure A.4: User Interface of Yieldwatch

In this figure, we provide a snapshot of user-interface environment of Yieldwatch, a 3rd-party information platform. For a current snapshot, see https://www.yieldwatch.net/.

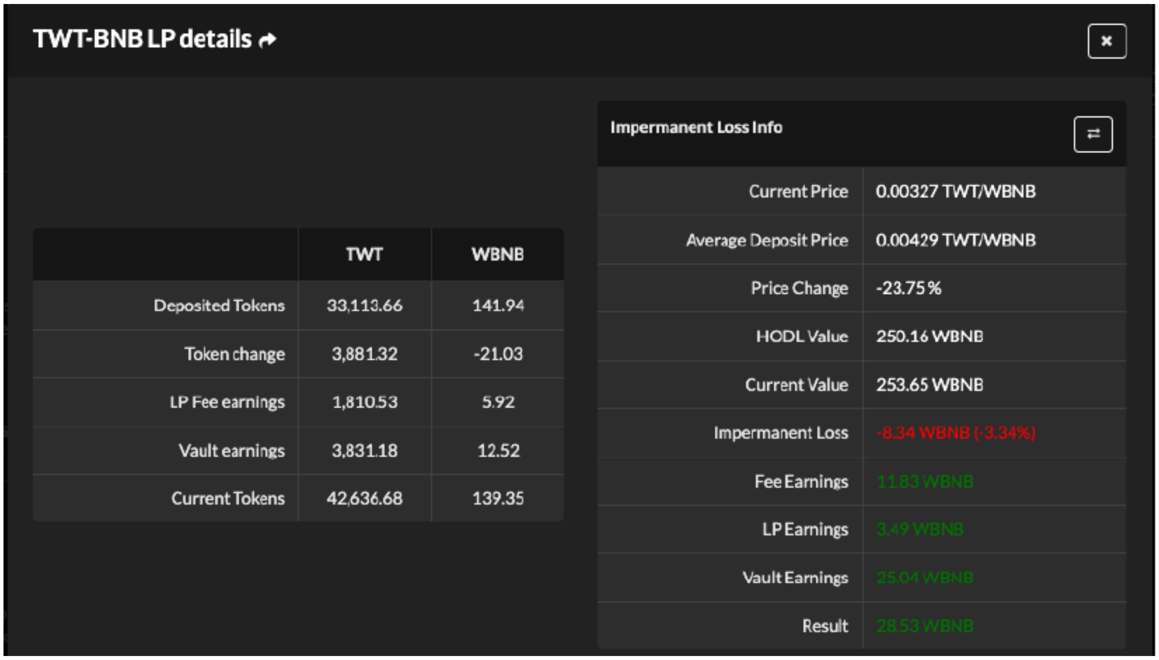


Figure A.5: The Impact of Price Divergence on Impermanent Loss

In this figure, we illustrate the impact of price divergence on impermanent loss, defined as the ratio of the portfolio value in the liquidity provision versus buy-and-hold strategies minus one (see Equation (B.2)). The y-axis indicates the impermanent loss (in %). The x-axis provides, for a representative pair of tokens A and B used for liquidity provision, a measure of price divergence over an h-period horizon defined as ρ_{t+h}/ρ_t, where $\rho_t = P_t^A/P_t^B$.

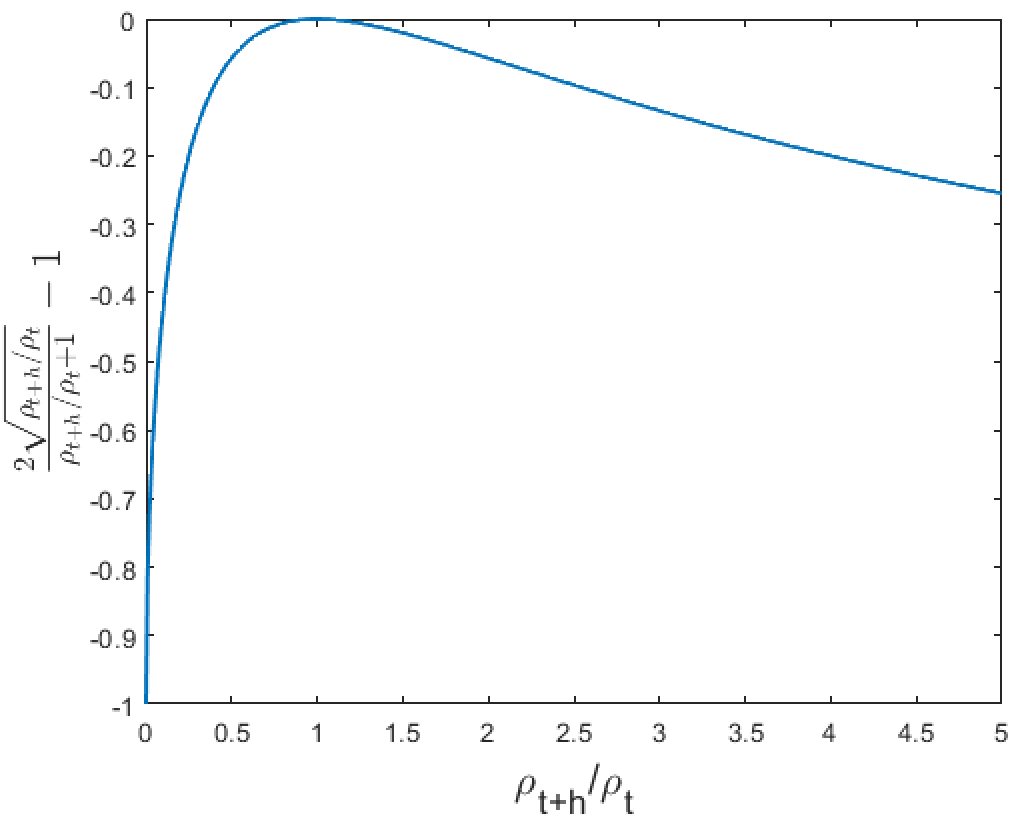
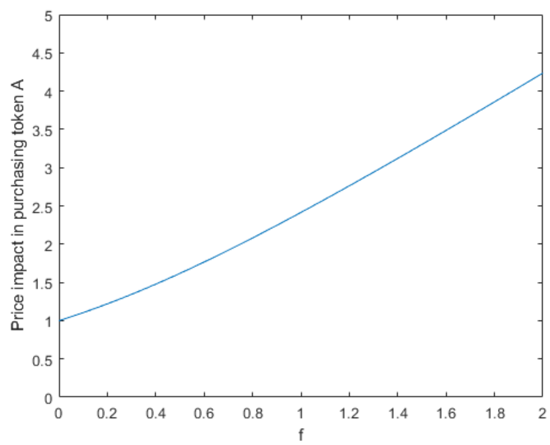


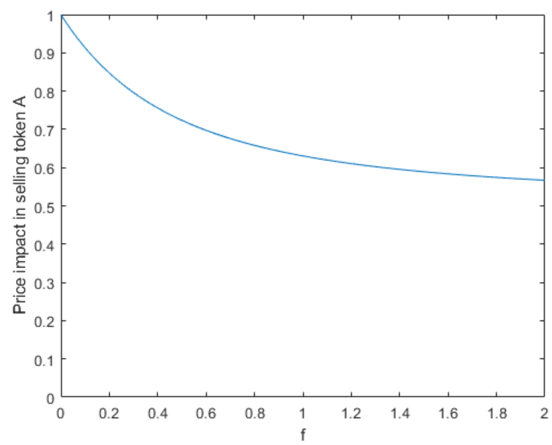
Figure A.6: Model-Implied Price Impact due to Yield Farming

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)

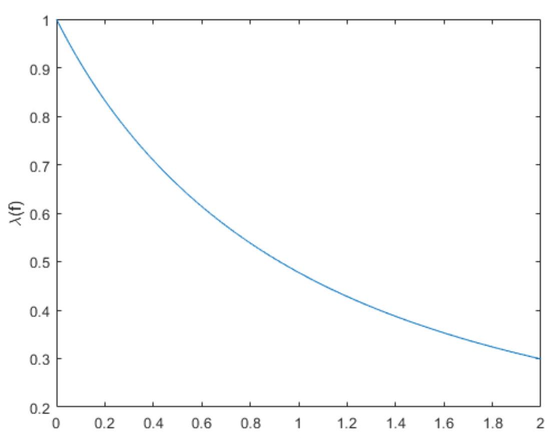


Figure A.7: Yield Farming Return Decomposition - Weekly Frequency

In this figure, we plot each component of weekly value-weighted returns across yield farms, by quintiles based on the magnitude of their total offered yields at the start of the week. The total offered farm yield is the sum of the offered yield and the trading fee yield estimated using the previous day's trading volume. Every week, we compute the weekly capital gain, impermanent loss, trading fee, and realized yield for all listed farms. Then, we take the average of each component across farms in each quintile using the size of each farm at the start of the week as weights. In Panels (a) to (d), the blue bars illustrate the average weekly capital gain, impermanent loss, trading fee, and realized yield. The red error bars plot their associated 95% confidence intervals. The mean of each component is displayed above their respective bars.

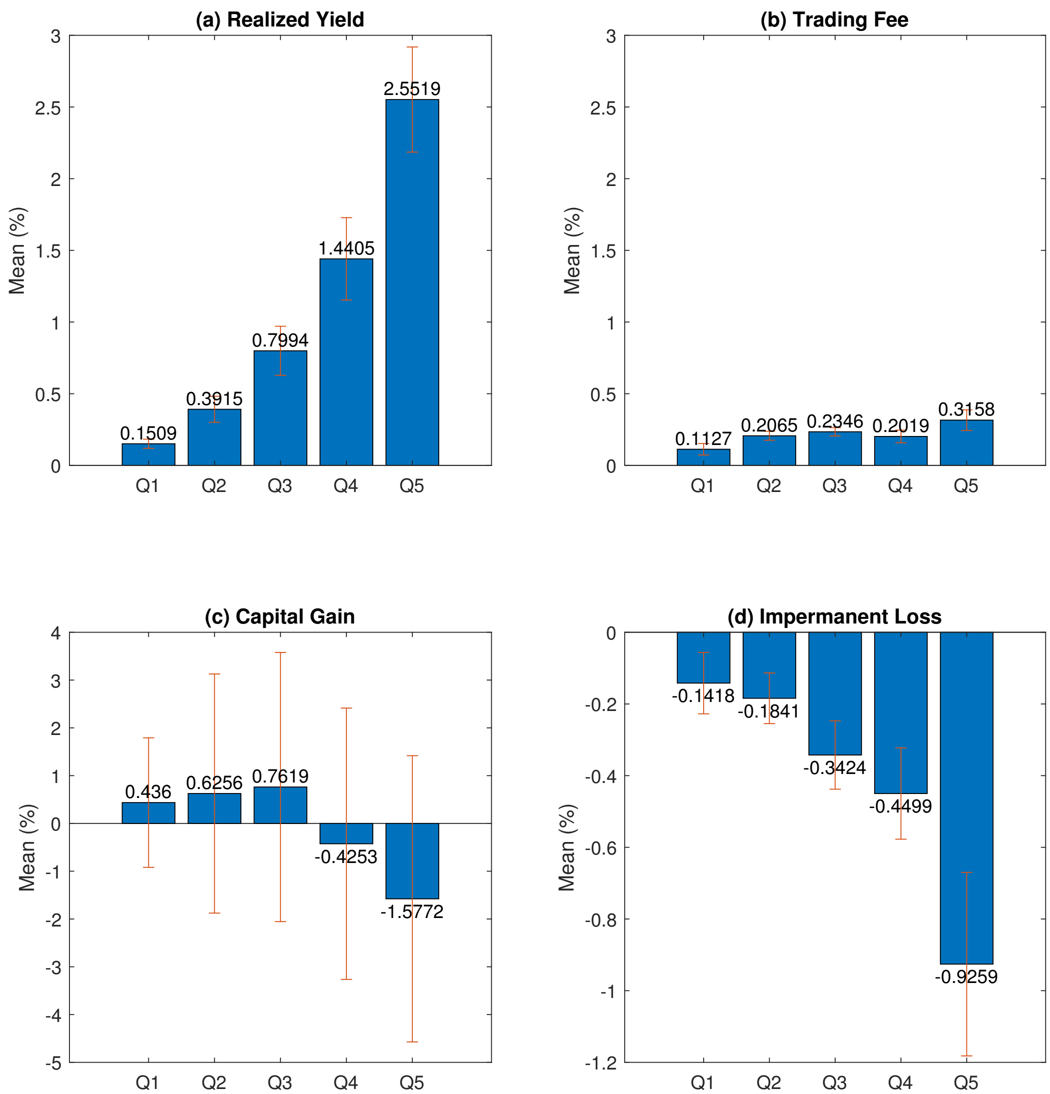


Figure A.8: Risk-Adjusted Returns from Yield Farming - Weekly Frequency

In this figure, we plot average risk-adjusted returns (i.e., alphas) and their associated 90% & 95% confidence intervals for different trading strategies at the weekly trading frequency. In Panel (a), we compare the performance of yield farming to that of liquidity mining without considering trading frictions. On each day, we sort farms into quintiles based on their in-sample total offered farm yields. The total offered farm yield is the sum of the offered yield and the trading fee yield estimated using the previous day's trading volume. In each quintile, we form value-weighted portfolios by using size of the liquidity pools as weights. A yield farming strategy is a strategy in which investors not only earn trading fee revenue but also farm yields, whereas investors that restrict themselves to liquidity mining can only earn trading fee revenue. We estimate alphas from a three factor model based on the work of Liu, Tsyvinski, and Wu (2022) and also account for the performance of BNB. The circle (square) and the associated bar display alphas and their 95% confidence intervals for yield farming (liquidity mining) without considering frictions. In Panel (b), we follow a similar procedure but provide alphas for yield farming strategies without trading frictions, yield farming strategies with frictions including gas fees, trading fees, and price impact, and yield farming strategies considering not only the frictions but also investor mistakes. We describe detailed trading strategies in Section 5.5.

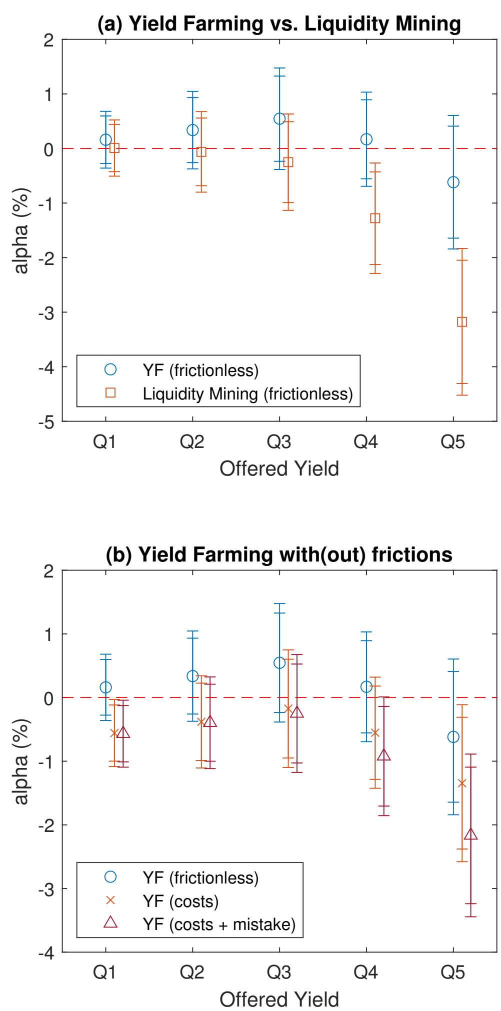
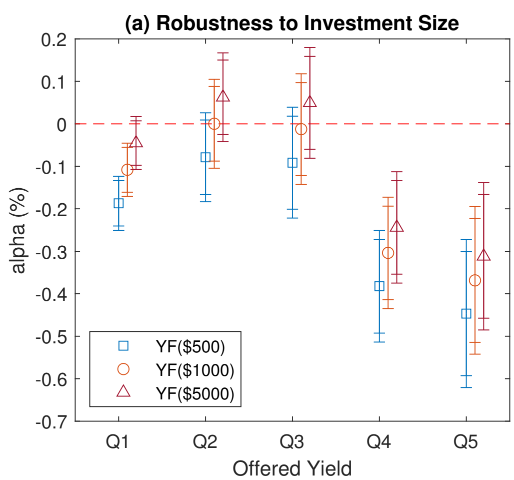


Figure A.9: Risk-Adjusted Returns from Yield Farming - Robustness

In this figure, we plot average risk-adjusted returns (i.e., alphas) and their associated 90% & 95% confidence intervals for different trading strategies, varying certain parameter choices. The starting set of parameters include an investment size of \$1,000, investment duration (i.e., time to rebalance) of 10 days, and diversification across two farms. In Panel (a), we change the duration, keeping all other parameters fixed. In Panel (b), we change the investment size keeping all other parameters fixed. We estimate alphas from a three-factor model based on the work of Liu, Tsyvinski, and Wu (2022) and also account for the performance of BNB. We account for frictions including gas fees, trading fees, price impact, and investor mistakes.



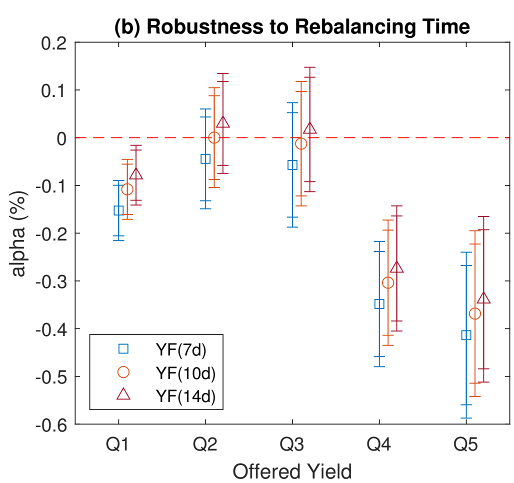


Figure A.10: The Impact of Aggregate Farm Multiplier Changes on Flows

In this figure, using an event study, we illustrate how changes in the aggregate CAKE allocation multiplier, ΔM_t, affect flows to farms. In Panel (a), we measure cumulative flows net of any price growth from one day before the event date $(t-1)$ to date $t+h$ by measuring the change in # LP tokens. In Panel (b), we measure the growth rate of the dollar value of the pool. We are interested in changes in cumulative flows and dollar value of the pools that are driven by shocks to the aggregate multiplier (M_t) due to changes to the multipliers of other farms, i.e., $\Delta m_{j,t} \mid j \neq i$, while $\Delta m_{i,t} = 0$. These shocks need to be large enough to have a meaningful impact on M_t and, therefore, $y_{i,t}$. We identify 4 events where $\Delta m_{i,t} = 0$ with $|\Delta M_t / M_t| > 0.15$. These 4 events are associated with increases in M_t. We then plot the average change in the outcome variables around the event dates, using a simple event study analysis. Specifically, we plot the coefficients β_k (and their 95% confidence intervals) from a regression $y_{i,t+h} = \alpha + \sum_{k=-7,k\neq-1}^{k=7} \beta_k I_k + Event \times FarmFE + \varepsilon_{i,t+h}$, where $y_{i,t+h}$ is defined as either $log(\frac{outstanding\ LP\ tokens_{i,t+h}}{outstanding\ LP\ tokens_{i,t-1}})$ or $log(\frac{\$\ of\ pool_{i,t+h}}{\$\ of\ pool_{i,t-1}})$. We cluster the standard errors at the farm and date levels.

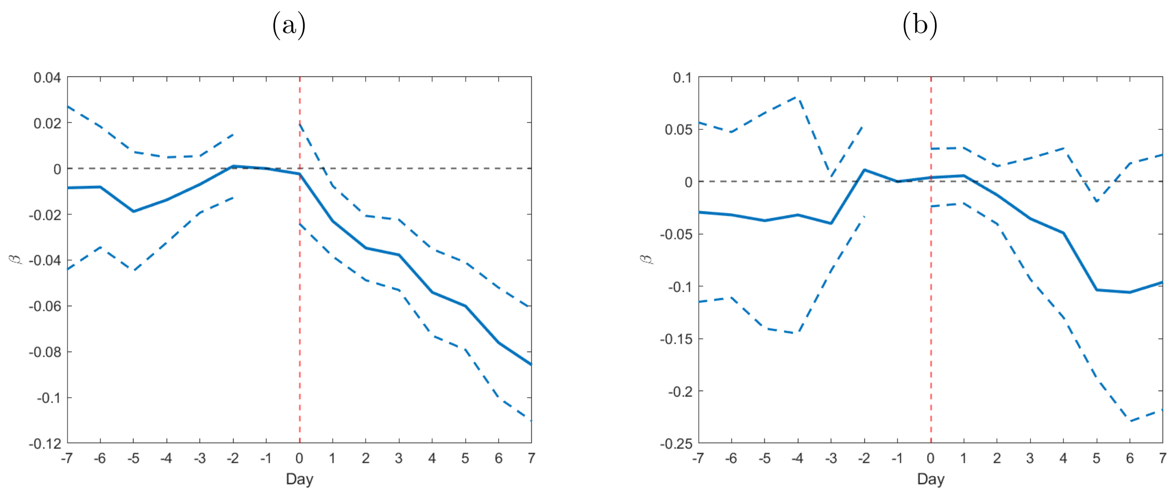
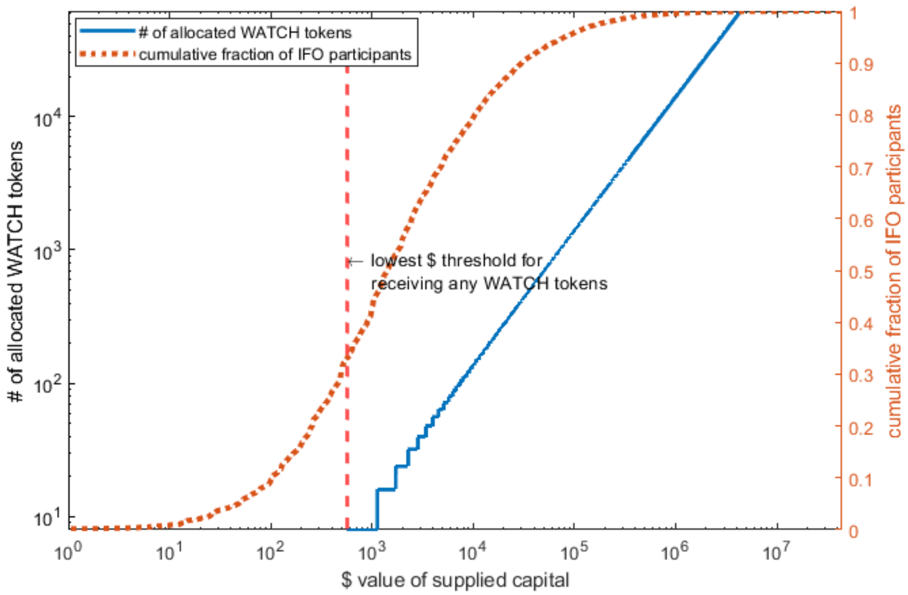


Figure A.11: YieldWatch Initial Farm Offering Allocation Schedule

Panel (a) plots the number of allocated WATCH tokens and the cumulative fraction of participants as a function of the dollar value of supplied capital (log scale) in the initial farm offering (IFO). Panel (b) illustrates the same relation between the allocated token amount/cumulative fraction of IFO participants and the aggregate bid amount, but restricts the x-axis to less than $1,200 USD. In both graphs, the dashed vertical line at $569.4 USD indicates the lowest dollar threshold for receiving any WATCH tokens. Bidders below that threshold did not receive any tokens. Source: PancakeSwap Yieldwatch IFO contract.

(a)



(b)

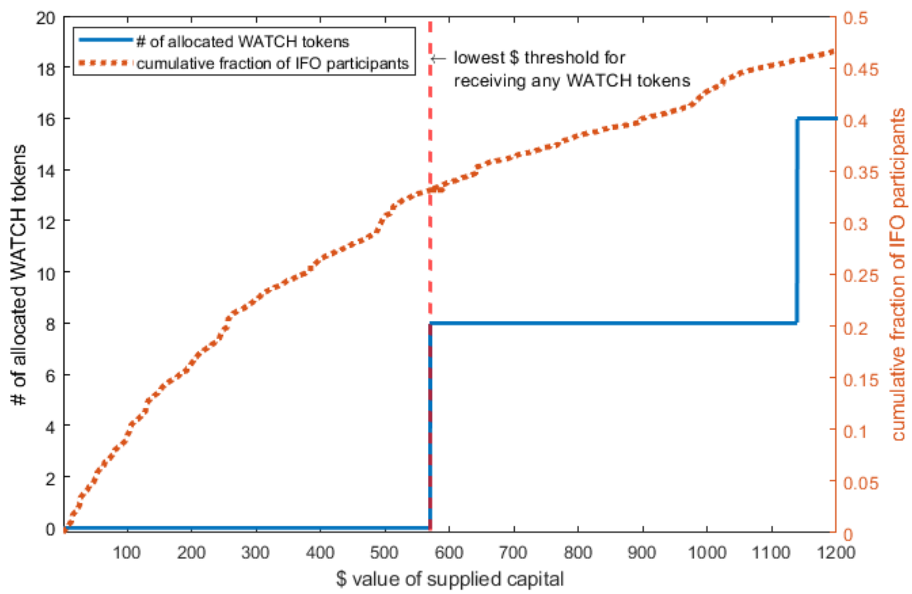


Figure A.12: YieldWatch Initial Farm Offering Capital Suply

This figure plots the time series of capital offered in the Yieldwatch initial farm offering (IFO) for the acquisition of WATCH tokens. The time scale indicates the number of seconds elapsed since the start of the IFO. Source: PancakeSwap Yieldwatch IFO contract.

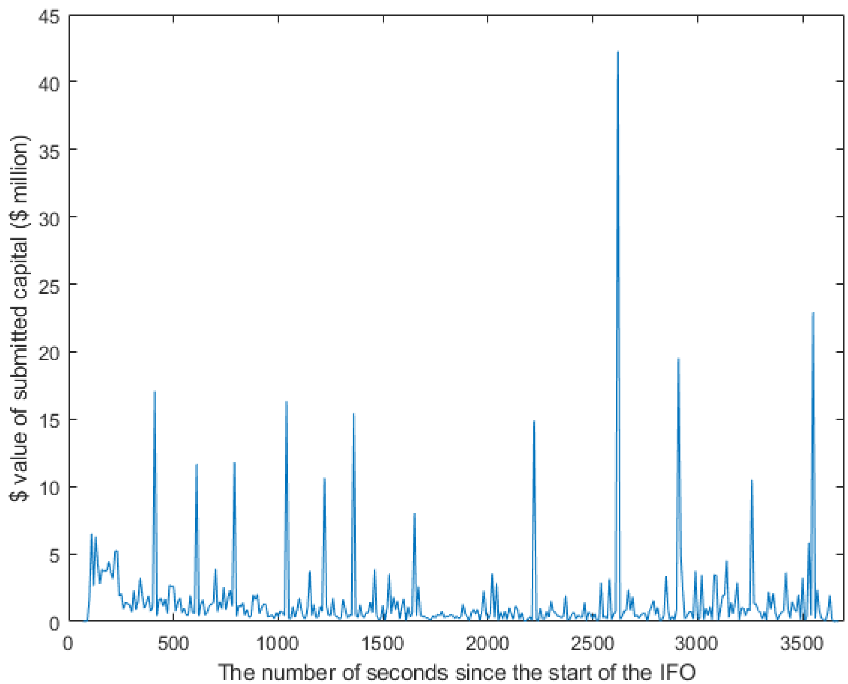


Figure A.13: YieldWatch DiD Parallel Trends

In this figure, we report the results from a difference-in-differences regression for the impact of Yieldwatch token acquisition on the reaching for yield propensity. Specifically, for farmer i in farm j at time t, we run the regression:

$$Flow_{t,t+7}^{i,j} = \alpha + \sum_{k=-3,k\neq -1}^{k=3} \beta_k I_k \times Yieldwatch^i \times Total\_Offered\_Yield_t^j + ... + \varepsilon_t^{i,j},$$

as per regression specifications reported in Table 8, where $Flow_{t,t+7}^{i,j}$ is the net inflow (token growth) into farm j by famer i, I_k are quarterly time indicator variables cast in event time around the Yieldwatch token acquisition, $Yieldwatch^i$ is one if a farmer ever holds Yieldwatch tokens or provides liquidity to the WATCH-BNB pool, and zero otherwise, and $Total\_Offered\_Yield_t^j$ is the total yield offered by farm j. The total offered farm yield is the sum of the offered yield and the trading fee yield estimated using the previous day's trading volume. We use the quarter just prior to the token acquisition as benchmark. Standard errors are clustered at the farmer level and by week. In the figure, we report 95% and 90% confidence bounds. The sample period is September 23, 2020 to July 31, 2022. The vertical dashed line indicates the quarter of the first Yieldwatch token acquisition.

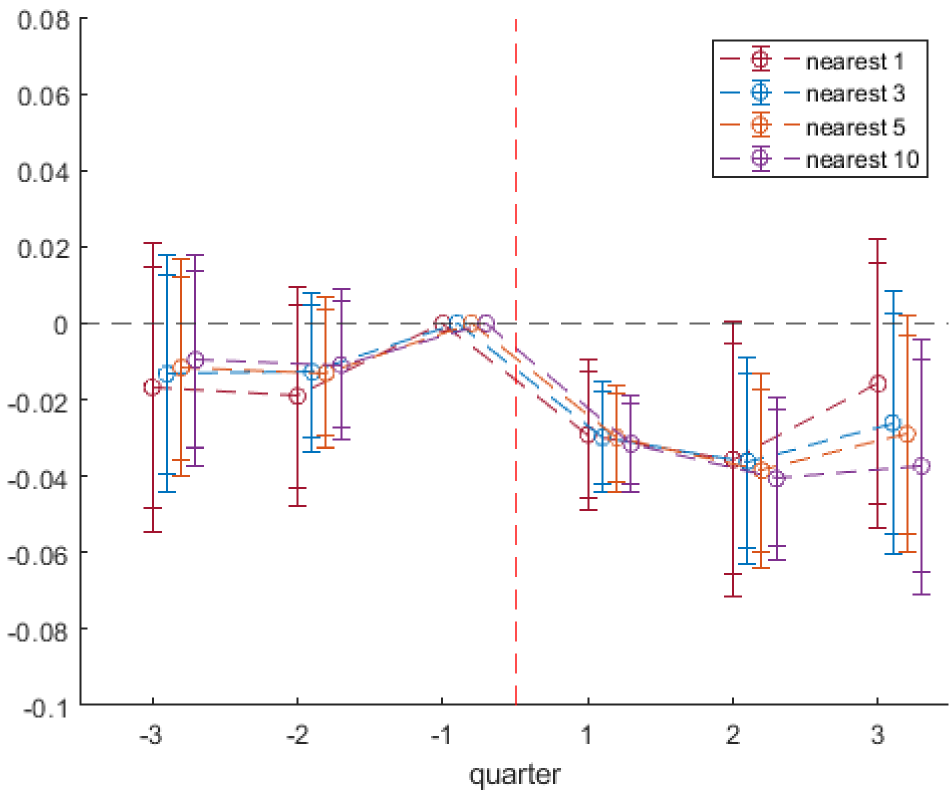


Table A.1: Literature on Decentralized Finance and Decentralized Exchanges

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 (2021)	✓		✓			✓		
Aoyagi (2021)	✓		✓			✓		
Aoyagi and Ito (2021)	✓		✓			✓	✓	
Neuder, Rao, Moroz, and Parkes (2021)	✓		✓			✓	✓	
Park (2023)	✓		✓			✓	✓	
Lehar and Parlour (2024)	✓	✓	✓			✓	✓	
Han, Huang, and Zhong (2021)		✓	✓				✓	
Capponi and Jia (2021)	✓	✓	✓	✓			✓	
Foley, O'Neill, and Putnins (2022)	✓	✓	✓	✓		✓	✓	
Fang (2023)		✓	✓			✓		
Li, Naik, Papanicolaou, and Schönleber (2024)		✓	✓			✓		
This study		✓			✓	✓		✓

Table A.2: 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.3: Top 10 Cryptocurrency Decentralized Exchanges

In this table, we report information about 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 spot assets 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 A.4: Determinants of Staking Ratios

In this table, we regress the staking ratio on a constant and indicator variables that take the value one if the staking ratio corresponds to the third farm ($3rd\ farm\ dummy$) to which the farmer provides liquidity and zero otherwise. Other indicator variables are defined similarly for the 4th ($4th\ farm\ dummy$), 5th ($5th\ farm\ dummy$), and more than five farms ($> 5th\ farm\ dummy$). The aggregate staking ratio is defined as the ratio of LP tokens staked in yield farms to the aggregate amount of LP tokens minted to certify liquidity provision. At the farmer/wallet level, we define the staking ratio to be the number of LP tokens staked to the main staking contract divided by the number of LP tokens received from PancakeSwap. The staking ratio is one if the investor's average staking ratio is greater than 0.99 and 0 otherwise. In columns (1) to (3), we deploy a linear probability model. In columns (4) to (6), we deploy a logistic regression model. Standard errors are clustered at the farmer level. *,**, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Staking Ratio (0 or 1)					
3rd farm dummy	0.1656***	0.0833***	0.0088***	0.8156***	0.9613***	0.2331***
	(0.0023)	(0.0018)	(0.0018)	(0.0135)	(0.0262)	(0.0505)
4th farm dummy	0.1982***	0.0991***	0.0133***	1.0247***	1.2614***	0.3711***
	(0.0029)	(0.0021)	(0.0018)	(0.0193)	(0.0369)	(0.0563)
5th farm dummy	0.2182***	0.1099***	0.0195***	1.1678***	1.5083***	0.5642***
	(0.0033)	(0.0021)	(0.0020)	(0.0244)	(0.0416)	(0.0657)
>5th farm dummy	0.2406***	0.1139***	0.0219***	1.3472***	1.6467***	0.6687***
	(0.0030)	(0.0024)	(0.0013)	(0.0251)	(0.0586)	(0.0390)
Constant	0.6239***			0.5062***		
	(0.0008)			(0.0034)		
Sample	All	All	inv. > $1000	All	All	inv. > $1000
Model		LPM			Logit	
Week FE	No	Yes	Yes	No	Yes	Yes
Farm FE	No	Yes	Yes	No	Yes	Yes
N	9,777,759	9,777,759	2,616,346	9,777,759	9,777,759	2,615,351
adj. R-sq	0.045	0.545	0.655			
pseudo R-sq				0.039	0.484	0.609

Table A.5: Comparison of Cryptocurrency Three-Factor Regressions

This table compares our regression results for portfolios sorted on one-week momentum by quintile to those reported in Liu, Tsyvinski, and Wu (2022). The sample period used in Liu, Tsyvinski, and Wu (2022) 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 as the period for our replication. We compare our parameter estimates to those obtained in Liu, Tsyvinski, and Wu (2022) (NBER version dated May 2019, strategy "r 1,0" in Table 11).

Panel A: Regressions from			**Quintile**		
Liu, Tsyvinski, and Wu (2022)	**1**	**2**	**3**	**4**	**5**
α	-0.015	-0.010	-0.003	0.025	-0.012
$t(\alpha)$	*-1.970*	*-1.535*	*-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

Table A.6: Summary Statistics of Coins used for Constructing Cryptocurrency Factors

In this table, we provide summary statistics of cryptocurrencies used for the construction of cryptocurrency factors as in Liu, Tsyvinski, and Wu (2022). Our sample period for cryptocurrency factors starts on December 28, 2013 and ends on July 31, 2022. 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	27	388.7469	7.3785	1.7534	0.0449
2014	110	249.6613	3.9564	1.1680	0.0304
2015	83	133.3719	2.6832	1.1360	0.0094
2016	177	157.6816	3.2750	1.7023	0.0183
2017	818	366.8033	8.8999	15.2997	0.1123
2018	1612	334.4740	8.9061	18.8046	0.1064
2019	1452	179.7373	5.4237	59.5064	0.1390
2020	1695	282.0723	6.7168	114.3506	0.2594
2021	2701	807.3235	15.0370	127.6496	0.5835
2022	2093	1,049.0538	14.2498	105.4045	0.5382

Table A.7: 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 frictions. We report all information from the perspective of an initial USD investment. We provide detailed description of parameters used to compute returns on each strategy in Section 4.5.1. On each day, we sort farms into quintiles based on their offered yields to form value-weighted portfolios. A yield farming strategy is a strategy in which investors not only earn trading fee revenue but also farm yields whereas in liquidity mining, investors can only earn trading fee revenue. *Frictionless Benchmark* (*Liquidity Mining*) refers to yield farming (liquidity mining) strategies with full staking and no frictions. *Yield Farming with Frictions* refers to yield farming strategies considering gas fee, trading fee, and price impact which adversely affect returns. *Yield Farming with Frictions & Investor Mistake* not only considers the frictions but also investors not fully staking in farms. Panel A (B) describes trading strategies where we rebalance portfolios every day (week). Excess returns are computed relative to the three-month U.S. Treasury bill secondary market rate from FRED. All returns are value-weighted using the pools' aggregate liquidity as weighing factors. The column (OBS) reports the number of observations. We report the mean return ($Mean$), the standard deviation, the Sharpe ratio (SR), information ratio (IR), the alpha from a three factor model based on the work of Liu, Tsyvinski, and Wu (2022) and BNB's performance, and the t-statistic for alpha from the three-factor regressions. The sample period is March 1, 2021 to July 31, 2022. Return-based statistics are not annualized.

Panel A: Daily							
Strategy	**Mean**	**SD**	**SR**	**IR**	α	**t-stat of α**	**OBS**
Yield Farming (Frictionless Benchmark)							
Quantile 1	0.0007	0.0242	0.0276	-0.0046	0.0000	-0.1038	518
Quantile 2	0.0022	0.0451	0.0492	0.0765	0.0011	1.9758	518
Quantile 3	0.0022	0.0465	0.0464	0.0681	0.0010	1.5542	518
Quantile 4	0.0000	0.0463	-0.0002	-0.1047	-0.0015	-2.2153	518
Quantile 5	-0.0001	0.0513	-0.0027	-0.0745	-0.0013	-1.5535	518
Liquidity Mining							
Quantile 1	0.0004	0.0242	0.0179	-0.0357	-0.0003	-0.7992	518
Quantile 2	0.0016	0.0450	0.0355	0.0335	0.0005	0.8720	518
Quantile 3	0.0009	0.0463	0.0203	-0.0090	-0.0001	-0.2044	518
Quantile 4	-0.0022	0.0460	-0.0478	-0.2505	-0.0036	-5.2177	518
Quantile 5	-0.0041	0.0509	-0.0806	-0.2860	-0.0052	-5.9415	518
Yield Farming with Frictions							
Quantile 1	-0.0004	0.0242	-0.0156	-0.1498	-0.0011	-3.3414	518
Quantile 2	0.0012	0.0451	0.0259	0.0012	0.0000	0.0307	518
Quantile 3	0.0011	0.0465	0.0238	-0.0005	0.0000	-0.0111	518
Quantile 4	-0.0011	0.0463	-0.0230	-0.1789	-0.0025	-3.7878	518
Quantile 5	-0.0012	0.0513	-0.0234	-0.1325	-0.0024	-2.7627	518
Yield Farming with Frictions & Investor Mistakes							
Quantile 1	-0.0004	0.0242	-0.0162	-0.1514	-0.0011	-3.3756	518
Quantile 2	0.0012	0.0451	0.0256	0.0001	0.0004	0.0032	518
Quantile 3	0.0010	0.0465	0.0213	-0.0082	-0.0001	-0.1877	518
Quantile 4	-0.0016	0.0463	-0.0349	-0.2147	-0.0030	-4.5383	518
Quantile 5	-0.0025	0.0512	-0.0495	-0.2031	-0.0037	-4.1603	518

Panel B: Weekly							
Strategy	**Mean**	**SD**	**SR**	**IR**	α	t**-stat of** α	**OBS**
Yield Farming (Frictionless Benchmark)							
Quantile 1	0.0056	0.0584	0.0955	0.0783	0.0016	0.6040	74
Quantile 2	0.0104	0.1089	0.0954	0.1136	0.0034	0.9304	74
Quantile 3	0.0145	0.1240	0.1172	0.1183	0.0055	1.1495	74
Quantile 4	0.0077	0.1236	0.0621	0.0489	0.0017	0.3849	74
Quantile 5	0.0036	0.1333	0.0273	-0.1200	-0.0062	-0.9904	74
Liquidity Mining							
Quantile 1	0.0041	0.0582	0.0699	0.0039	0.0001	0.0307	74
Quantile 2	0.0065	0.1082	0.0599	-0.0212	-0.0006	-0.1661	74
Quantile 3	0.0065	0.1224	0.0534	-0.0554	-0.0025	-0.5555	74
Quantile 4	-0.0067	0.1224	-0.0550	-0.3366	-0.0128	-2.4740	74
Quantile 5	-0.0219	0.1304	-0.1678	-0.6093	-0.0318	-4.6301	74
Yield Farming with Frictions							
Quantile 1	-0.0017	0.0582	-0.0288	-0.2719	-0.0056	-2.0765	74
Quantile 2	0.0031	0.1087	0.0288	-0.1270	-0.0038	-1.0333	74
Quantile 3	0.0072	0.1237	0.0586	-0.0379	-0.0017	-0.3698	74
Quantile 4	0.0004	0.1234	0.0030	-0.1585	-0.0055	-1.2409	74
Quantile 5	-0.0037	0.1330	-0.0279	-0.2610	-0.0135	-2.1383	74
Yield Farming with Frictions & Investor Mistakes							
Quantile 1	-0.0018	0.0582	-0.0307	-0.2774	-0.0057	-2.1175	74
Quantile 2	0.0030	0.1087	0.0275	-0.1322	-0.0040	-1.0764	74
Quantile 3	0.0065	0.1240	0.0527	-0.0541	-0.0025	-0.5301	74
Quantile 4	-0.0034	0.1242	-0.0270	-0.2504	-0.0092	-1.9381	74
Quantile 5	-0.0123	0.1322	-0.0927	-0.4188	-0.0217	-3.3233	74

Table A.8: APY.Vision NFT Airdrops

This table describes the 20 airdrops for APY.Vision NFTs for which we could identify useful information on eligibility for receiving NFTs. The column *Airdrop* lists the name of an airdrop. The column *Eligibility requirements* describes the eligibility requirements for receiving NFTs. The column *Procedure of collecting data* explains how we collect data to identify wallets eligible for receiving NFTs. The column *# NFT receivers/Receipt dates* lists the number of NFT receivers and the starting time of the corresponding airdrop. The column *Matched # NFT receivers/# eligible wallets* lists the number of wallets that received the NFTs and the number of wallets that were eligible to receive NFTs by satisfying the eligibility requirements.

Airdrop	**Eligibility requirements**	**Procedure for collecting data**	**# NFT receivers/Receipt dates**	**Matched # NFT receivers/# eligible wallets**
Balancer + APY Vision	Anyone holding veBAL on April 7, 2022, 10 PM UTC.	**Approach** • 1. Download the transfer data for 80BAL20WETH LP tokens. • 2. Find the wallets that staked the 80BAL20WETH LP to the veBAL contract address as of the date for a snapshot. **Adjustment** • We find that 23/30 are matched if we use "April 7, 10 PM UTC" as the date for snapshot. If we use April 8, 9 PM instead, 30/30 are matched. So we decided to use April 8, 9 PM instead.	30/Apr-08-2022 09:36:11 PM	30/439
APY Vision X YAxis 2022	Provide liquidity to the YAXIS/ETH pool on yAxis between the dates Feb 14 and Feb 28. Users who already have funds in the vaults are already eligible.	**Approach** • 1. Download the transfer data for YAXIS/ETH LP tokens. • 2. Find wallets that have ever provided liquidity to the pool as of Feb 28, 2021, 11:59:59 PM.	29/Mar-07-2022 10:23:15 PM	29/595
Tesseract X APY Vision	Provide liquidity to one of the vaults on Tesseract between the dates Dec 17 and Dec 24. Users who already have funds in the vaults are already eligible. Eligible vaults: WMATIC, DAI, USDC, WETH, WBTC.	**Approach** • 1. Download the tVault tokens for the pools. • 2. For each tVault token, find wallets that have ever held tv tokens as of 11:59:59 PM on Dec. 24, 2021.	9/Dec-29-2021 02:05:21 AM	9/3598

Continued on next page

Table A.8 – *Continued from previous page*

Airdrop	Eligibility requirements	Procedure for collecting data	# NFT receivers/Receipt dates	Matched # NFT receivers/# eligible wallets
Swapr + APY Vision	To be eligible to win the NFT pictured above, provide liquidity to the SWPR/wETH pool on Swapr between the dates Nov 8 and Nov 15 (A new position is not required existing LPs are eligible)	**Approach** • 1. Download the SWPR-wETH LP token transfers data. • 2. Find wallets that have ever provided liquidity to the SWPR-wETH pool as of 11:59:59 PM on Nov. 15, 2021. **Additional Information** • To identify wallets that have ever provided liquidity, we choose wallets that ever received the LP tokens as a consequence of execution of methods starting with "addLiq."	15/Nov-17-2021 04:45:56 AM	15/851
Trader Joe x APY Vision	Provide liquidity to the JOE-AVAX pool on Trader Joe between the dates Oct 26 and Nov 2.	**Approach** • 1. Download JOE-AVAX LP token transfers data from Avalanche SnowTrace. • 2. Find wallets that have ever provided liquidity to the pool as of 11:59:59 PM on Nov. 2, 2021.	15/Nov-03-2021 10:30:58 PM	15/12944
CoinGecko + APY Vision #1	(1) Purchase CoinGecko Candy Code Voucher. (2) Enter your Candy Voucher Code in the input field. (3) Follow CoinGecko on Twitter. (4) Follow APY.vision on Twitter. (5) Join APY.vision on Discord. (6) Retweet This Tweet.	**Approach** • We need to know who are CoinGecko Candy holders but CoinGecko Candies are not cryptocurrencies, and therefore, we cannot get this information. So, we cannot perfectly do this.	12/Oct-26-2021 09:16:56 PM	12/N.A.
CoinGecko + APY Vision #2	(1) Purchase CoinGecko Candy Code Voucher. (2) Enter your Candy Voucher Code in the input field. (3) Follow CoinGecko on Twitter. (4) Follow APY.vision on Twitter. (5) Join APY.vision on Discord. (6) Retweet This Tweet.	**Approach** • We need to know who are CoinGecko Candy holders but CoinGecko Candies are not cryptocurrencies, and therefore, we cannot get this information. So, we cannot perfectly do this.	232/Oct-26-2021 09:11:20 PM	232/N.A.
xDai + APY Vision	Provide delegator services for the xDai chain between the dates Oct 5 and Oct 12. The lucky winners will receive an XDAIxAPY NFT that allows access to the PRO edition of APY.vision until 01/01/2022. A total of 10 NFTs will be given to the winners.	**Approach** • 1. Download transfer data for STAKE token. • 2. Find wallets that have ever staked the STAKE tokens to the staking contract as of 11:59:59 PM on Oct. 12, 2021. **Additional Information** • We find that all NFT receivers sent their STAKE tokens to the staking contract before the NFT airdrop. Therefore, we use this approach to identify wallets that provide delegator services.	10/Oct-18-2021 11:45:22 PM	10/521
APY Vision V2 Release Party NFT	No information available.	No information available.	(1) 14/Jul-22-2021 03:48:35 AM (2) 12/Jul-23-2021 05:27:15 PM	26/N.A.

Continued on next page

Table A.8 – *Continued from previous page*

Airdrop	Eligibility requirements	Procedure for collecting data	# NFT receivers/Receipt dates	Matched # NFT receivers/# eligible wallets
Lift Kitchen x APY Vision NFT #1	To enter, you must have over $10K USD deposited in either the lfBTC-LIFT or the wBTC-lfBTC pool.	**Approach** • 1. Download transfer data for WBTC-LFBTC and LIFT-LFBTC tokens. • 2. Find wallets that have provided liquidity to those pools as of 8:00:00 PM on July 2, 2021, because the first NFT receipt date was Jul-02-2021 08:21:24 PM. **Additional Information** • If we use only WBTC-LFBTC and LIFT-LFBTC, we can only find 15 matches out of 20. If we add ETH-LFBTC, we can find 20/20.	20/Jul-02-2021 08:21:24 PM	20/220
YAxis + APY Vision Vault NFT	How do I enter? You must stake your yAxis tokens here! (← the staking contract)	**Approach** • 1. Download transfer data for yAxis. • 2. Find wallets that have ever staked yAxis to the staking contract as of – (1) 9:00:00 PM on May 25, 2021. (The first NFT receipt date was May 25, 2021, 09:36:11 PM.) – (2) 10:00:00 PM on June 1, 2021. (The last NFT receipt date was June 1, 2021, 10:28:17 PM.)	(1) 5/May-25-2021 09:36:11 PM (2) 5/Jun-01-2021 10:28:17 PM	10/4482
APY Vision Ambassador NFT	No information available.	No information available.	15/Jul-22-2021 03:48:35 AM	15/N.A.
Index Coop x APY Vision Owl	Supply over $10K liquidity to the new MVI/ETH pool on the first day - You can win by entering the MVI/ETH pool with over $10,000 of liquidity in the first 24 hours.	**Approach** • 1. Download transfer data for MVI-ETH LP tokens. • 2. Find wallets that have ever provided liquidity to the pool as of – (1) 4:00:00 PM on April 9, 2021. (The first NFT receipt date was Apr-09-2021 04:15:33 PM.) – (2) 5:00:00 PM on April 26, 2021. (The first NFT receipt date was Apr-26-2021 05:35:50 PM.) **Additional Information** • I used 4/9/2021 noon as the time for a snapshot. There are two batches: April 9 and April 26.	(1) 10/Apr-09-2021 04:15:33 PM (2) 13/Apr-26-2021 05:35:50 PM	12/509

Continued on next page

Table A.8 – *Continued from previous page*

Airdrop	Eligibility requirements	Procedure for collecting data	# NFT receivers/Receipt dates	Matched # NFT receivers/# eligible wallets
Pickle x APY.Vision NFT	You must provide liquidity to the PICKLE/ETH Uniswap V2 Pools and then stake that in the farm.	**Approach** • 1. Download transfer data for PICKLE-ETH LP tokens. • 2. Find wallets that have ever staked the LP tokens to the staking contract as of – (1) 10:00:00 PM on March 10, 2021. (The first NFT receipt date was Mar-10-2021 10:32:07 PM.) – (2) 10:00:00 PM on March 17, 2021. (The first NFT receipt date was Mar-17-2021 10:15:32 PM.) – (3) 11:00;00 PM on March 24, 2021. (The first NFT receipt date was Mar-24-2021 11:51:39 PM.)	(1) 5/Mar-10-2021 10:32:07 PM (2) 5/Mar-17-2021 10:15:32 PM (3) 5/Mar-24-2021 11:51:39 PM	15/13295
Benchmark Protocol x APY.Vision NFT	You must provide liquidity to the MARK/ETH or MARK/USDC Uniswap V2 Pools and then stake that pool on "The Press" at Benchmark Protocol.	**Approach** • 1. Download transfer data for MARK-ETH, MARK-USDC LP tokens. • 2. Fine wallets that have ever staked the LP tokens to the Benchmark Protocol's staking contract as of – (1) midnight on March 6, 2021. (The first NFT receipt date was Mar-06-2021 12:16:15 AM.) – (2) 7:00:00 PM on March 13, 2021. (The first NFT receipt date was Mar-13-2021 07:58:48 PM.)	(1) 5/Mar-06-2021 12:16:15 AM (2) 5/Mar-13-2021 07:58:48 PM	10/1079
Alpha Finance 2.0 x APY.Vision NFT	On Monday, starting from Feb 8th, we will randomly select and announce 5 addresses on Twitter that will receive these NFTs. The 5 addresses will be randomly selected from a pool of addresses that use Alpha Homora V2 in the previous week (in this case from Feb 1st to Feb 8th). The program will run for 4 weeks, and in total, 20 NFTs will be distributed.	**Approach** • 1. Download all transactions executing functions of Alpha Router. • 2. Find wallets that have ever executed any functions of Alpha Router as of – (1) 8:00:00 PM on February 8, 2021. (The first NFT receipt was Feb-08-2021 08:47:39 PM.) – (2) 7:00:00 AM on February 20, 2021. (The first NFT receipt was Feb-20-2021 07:37:25 AM.) – (3) 11:00:00 PM on March 5, 2021. (The first NFT receipt was Mar-05-2021 11:56:43 PM.)	(1) 5/Feb-08-2021 08:47:39 PM (2) 10/Feb-20-2021 07:37:25 AM (3) 5/Mar-05-2021 11:56:43 PM	20/1843

Continued on next page

Table A.8 – *Continued from previous page*

Airdrop	Eligibility requirements	Procedure for collecting data	# NFT receivers/Receipt dates	Matched # NFT receivers/# eligible wallets
88MPHxAPY.Vision NFT	If you already stake in the MPH-ETH farm contract, you're already eligible to participate in the giveaway! If you haven't staked yet, you'll need to stake the Uniswap MPH-ETH liquidity pool (UNI-V2) tokens in the 88MPH staking contract, located at https://88mph.app/farming.	**Approach** • 1. Download transfer data for MPH-ETH LP tokens. • 2. Find wallets that ever staked the LP tokens in the staking contract as of – (1) 11:00:00 PM on December 27, 2020. (The first NFT receipt was Dec-27-2020 11:49:31 PM.) – (2) 6:00:00 PM on January 6, 2021. (The first NFT receipt was Jan-06-2021 06:57:58 PM.) – (3) midnight January 19, 2021. (The first NFT receipt was Jan-19-2021 12:46:38 AM.)	(1) 10/Dec-27-2020 11:49:31 PM (2) 10/Jan-06-2021 06:57:58 PM (3) 5/Jan-19-2021 12:46:38 AM	21/2946
SUSHISWAPxAPY.Vision NFT	If you already staked in the SUSHI-ETH farm contract, you're already eligible to participate in the giveaway! If you haven't staked yet, you'll need to stake the SushiSwap SUSHI-ETH liquidity pool (UNI-V2) tokens in the SushiSwap staking contract, located at https://sushiswap.fi (min. 1 ETH liquidity value is required to participate). We will be randomly drawing 25 winners from the list of addresses that have staked in the farming contract between now and 23:59:59 Dec 23, 2020 (GMT).	**Approach** • 1. Download transfer data for SUSHI-ETH LP tokens. • 2. Find wallets that ever staked the LP tokens in the SushiSwap staking contract as of Jan-06-2021 09:00:00 PM because the first NFT receipt date was Jan-06-2021 09:40:50 PM. **Adjustement** • If we use Dec 23, 2020, we cannot find that the treated wallets have SUSHI-ETH LP. So we used the later snapshot, January 6, 2021, 4 PM. Then, we tracked the wallets of those who had ever staked until then.	10/Jan-06-2021 09:40:06 PM	10/6511
INDEXxAPY.Vision NFT	Stake - If you already stake in the DPI/ETH farm contract, you're already eligible to participate in the giveaway! If you haven't staked yet, you'll need to stake the Uniswap DPI/ETH liquidity pool (UNI-V2) tokens in the INDEX staking contract, located at https://www.indexcoop.com/farm.	**Approach** • 1. Download transfer data for DPI-ETH LP tokens. • 2. Find wallets that ever staked the LP tokens in the INDEX staking contract as of 6:00:00 PM on Dec.22, 2020, because the first NFT receipt date was Dec 22, 2020, 06:04:27 PM.	10/Dec 22, 2020, 06:04:27 PM	10/855

Continued on next page

Table A.8 – *Continued from previous page*

Airdrop	Eligibility requirements	Procedure for collecting data	# NFT receivers/Receipt dates	Matched # NFT receivers/# eligible wallets
Harvest + APY.Vision Partnership	If you already staked in the UNISWAP-LP (FARM/USDC LPs) farm on Harvest.finance, you're already eligible to participate in the giveaway! If you haven't staked yet, you'll need to stake the eligible Uniswap FARM/USDC liquidity pool (UNI-V2) tokens in the Harvest.finance Farm.	**Approach** • 1. Download transfer data for FARM-USDC and FARM-ETH LP tokens. • 2. Find wallets that have ever staked the LP tokens in the Harvest Finance staking contract as of – (1) 4:00:00 pm on December 12, 2020. (The first NFT receipt date was Dec-12-2020 04:34:50 PM.) – (2) 1:00:00 am on December 20, 2020. (The first NFT receipt date was Dec-20-2020 01:42:32 AM.) – (3) 11:00:00 pm on December 29, 2020. (The first NFT receipt date was Dec-29-2020 11:58:26 PM.) – (4) 6:00:00 pm on January 12, 2021. (The first NFT receipt date was Jan-12-2021 06:21:09 PM.) – (5) 2:00:00 am on January 24, 2021. (The first NFT receipt date was Jan-24-2021 02:19:11 AM.) – (6) 5:00:00 am on February 10, 2021. (The first NFT receipt date was Feb-10-2021 05:15:42 AM.) – (7) midnight on March 2, 2021. (The first NFT receipt date was Mar-02-2021 12:18:39 AM.) **Adjustment** • We find that 67/70 are matched using this approach. We find that many others staked their FARM-ETH LP tokens instead of FARM-USDC LP tokens. Therefore, we added FARM-ETH LP tokens to find the eligible wallets.	(1) 10/Dec-12-2020 04:34:50 PM (2) 10/Dec-20-2020 01:42:32 AM (3) 10/Dec-29-2020 11:58:26 PM (4) 10/Jan-12-2021 06:21:09 PM (5) 10/Jan-24-2021 02:19:11 AM (6) 10/Feb-10-2021 05:15:42 AM (7) 10/Mar-02-2021 12:18:39 AM	70/17195