The Tokenomics of Staking*

Lin William Cong† Zhiheng He§ Ke Tang¶

First draft: Dec. 2020; this draft: Oct. 2023.

Abstract

Tokens offer transaction convenience in digital networks and are frequently staked for base-layer consensus generation or incentivizing platform development (consequently earning stakers rewards). We build a continuous-time model where heterogeneous agents dynamically allocate their wealth over staking, onchain transaction, and off-chain consumption. In this mean-field game with stochastic control and systematic shocks, aggregate staking ratios are key in linking staking to token pricing. Empirical findings from stakable tokens corroborate the model predictions: (i) staking ratio positive correlates with reward rates; (ii) higher staking ratios predict greater excess returns; (iii) uncovered interest rate parity is violated, generating significant crypto carry premia.

Keywords: Carry, Cryptocurrencies, DeFi, Heterogeneity, Inequality, PoS, Redistribution, UIP.

*We thank Bruno Biais, Philip Bond, Jonathan Chiu, Rod Garratt, Niklas Haeusle, Shiyang Huang, Huisu Jang, Stephen Karolyi, Leonid Kogan, Tao Li, Evgeny Lyandres, Urban Jermann, Julien Prat, Daniel Rabetti, Qihong Ruan, Fahad Saleh, Ville Savolainen, Donghwa Shin, Jingjie Zhang, Shunming Zhang, and conference and seminar participants at the Asian Bureau of Finance and Economic Research (ABFER) 10th Annual Conference, the 36th Australasian Finance and Banking Conference (AFBC), the 18th Annual Conference of the Asia-Pacific Association of Derivatives (APAD), the 19th Chinese Finance Annual Meeting (CFAM 2022), University of Cincinnati, Columbia Business School and School of Engineering and Applied Science (Digital Finance Seminar),Crypto and Blockchain Research Economics (CBER), the Department of the Treasury Office of Financial Research, the Economic Club of Memphis, Luohan Academy Webinar Series, Rizzo Center Decentralized Finance (DeFi) Conference at UNC Chapel Hill, the Fields-CFI Workshop on Mathematical Finance and Cryptocurrencies, FinTech Conference at the Fubon Center for Technology, Business and Innovation (NYU Stern), Halle Institute for Economic Research (IWH), HEC Paris, 2022 Hong Kong Conference for Fintech, AI, and Big Data in Business, Office of the Comptroller of the Currency (OCC) Novel Charters Working Group, The Pennsylvania State University Smeal College of Business, 4th Shanghai Financial Forefront Symposium, Tsinghua University PBC School of Finance, University of International Business and Economics School of Banking and Finance, the China Meeting of Econometric Society (CMES 2021, Shanghai), 7th PHBS Workshop in Macroeconomics and Finance, Ripple Labs London Onsite (Markets Team), Wolfe QES Virtual Global Innovation Conference, and Zhejiang University QingYuan Academy Digital Social Science Public Lecture Series for helpful comments and discussions. We are also grateful to the stakingrewards.com team for generously sharing the updated data that cover Sept 2020 - Feb 2022 for academic research and to the Fintech Dauphine Chair in partnership with Mazars and Crédit Agricole CIB, the Hogeg Blockchain Research Institute, as well as Ripple's University Blockchain Research Initiative (UBRI) for financial support. The paper subsumes key results from the working paper titled "Staking, Token Pricing, and Crypto Carry." Send correspondence to will.cong@cornell.edu.

†Cornell University and NBER. E-mail: will.cong@cornell.edu.

§Tsinghua University. E-mail: hezh19@mails.tsinghua.edu.cn.

¶Tsinghua University. E-mail: ketang@tsinghua.edu.cn.

1 Introduction

The past decade witnessed an explosive growth in blockchain-based platforms and cryptocurrencies, which totaled 2.5 trillion USD in market capitalization by the end of 2021, and a rising interest in Decentralized Finance (Harvey et al., 2021), which entails over 130 billion USD worth of assets as of Feb 2022. Platform tokens derive value by enabling users to complete economic transactions and hold stakes in the ecosystem, making them a hybrid of money and investible assets. The recent prevalence of token staking (value locking and yield farming, see, e.g., Augustin et al., 2022) for higher-layer DeFi innovations as well as for base layer consensus formation (e.g., through Proof-of-Stake, PoS, as discussed in John et al., 2022) further calls for a unified framework to understand the use of tokens as transaction media, investment assets, and deposit/collateral-like instruments. To this end, we offer the first study relating various functions tokens provide (e.g., transaction convenience and financial rewards through holding and staking) to token prices, both theoretically and empirically, with endogenous adoption and agent heterogeneity.

Specifically, we build a continuous-time model of an economy with a tokenized digital network, where agents optimally conduct transactions on a platform subject to demand shocks, stake tokens to earn rewards from both newly minted tokens and fees, and consume offline (off-chain or off-network). We cast the interactions as a mean-field game (MFG) with individual stochastic control. We incorporate systematic shocks and innovate upon the master equation approach, which marks a novel application of the methodology beyond macroeconomics and idiosyncratic shocks. The aggregate staking ratio is a key determinant for the equilibrium reward rate and token prices. The agents' onchain and overall wealth distributions in an expected steady state follow Pareto-like distributions, and staking facilitates redistribution from poorer or more usage-focused adopters to the richer or investment-focused adopters, as in an inflation tax. The wedge between usage convenience and financial returns explains the violation of uncovered interest rate parity (UIP) and predicts profitable carry trades.

Empirical evidence from a dataset covering all major stakable tokens also corroborates the model predictions. In particular, staking ratio is positively correlated with reward rates in the cross-section and negatively correlated in the time series. Moreover, we document that higher reward rates and wealth concentration attract greater future staking, increasing the aggregate staking ratio, which in turn positively predicts token excess returns.

To our knowledge, our study is also the earliest and most comprehensive investigation into the violations of uncovered interest rate parity and carry premia for stakable tokens, which dominate the cryptocurrency market (after Ethereum's move to PoS).

We conceptually highlight that staking involves two broad categories of activities: those related to pan-PoS consensus protocols and those in higher-layer (decentralized) applications.[1] Fundamentally, blockchain functions to generate a relatively decentralized consensus record of system states to enable economic interactions such as value or information exchanges (e.g., Cong and He, 2019). PoS consensus protocols have gained popularity with major market players such as Ethereum adopting them. Under PoS, agents stake native tokens to compete for the opportunity to record transactions, execute smart contracts, append blocks, etc., to earn block rewards and fees (e.g., Saleh, 2021; Kogan et al., 2021; Jermann, 2023). Meanwhile, various staking programs have become popular means for incentivizing desirable behavior in higher layer applications, escrowing a balance of tokens under custody in a smart contract or deploying them to enable network economic functionalities with stakers earning staking rewards (e.g., Augustin et al., 2022).

Our model applies to both categories and captures several distinguishing features of PoS protocols and stakable projects. First, such tokens are used on platforms that support specific economic transactions or broader use in onchain-based projects. This generates utility flows in, e.g., transaction convenience discussed in (Cong et al., 2021b; Biais et al., 2020). Second, the rate of staking rewards that an agent earns is influenced by other agents' behavior in aggregate, but individuals take it as given when making decisions.[2] Third, staking participation can influence the platform's development, e.g. by improving the efficiency and security of services, which enriches the agents' roles within the platform, although it is not a necessary feature in all the projects.

In our model, agents potentially heterogeneous in wealth and preference for accessing the digital network allocate and adjust their holdings of staked tokens, transaction tokens, and consumption numéraire under budget constraints, trading off staking rewards, token transaction convenience, and numéraire convenience for offline (off-chain or off-platform)

[1] The two are not mutually exclusive. Solana, for example, uses both PoS and DeFi staking. The classification we use follows mainstream cryptocurrency data aggregators such as CoinMarketCap. Even on non-blockchain-based or centralized platforms, various programs that involve escrows or crowd funds can be analyzed as a form of business layer staking through the lens of our framework.

[2] Polkadot (DOT) constitutes an example: the reward rate for validators is determined by the current aggregate staking ratio. The fewer DOTs are staked, the higher the yield is for a planned amount of reward.

consumption. Transaction convenience endogenously increases in platform productivity, which evolves over time and is potentially influenced by staking activities while the staking reward rate is jointly determined by aggregate reward and tokens staked.

In equilibrium, the staking reward rate solves a fixed point problem. Token price dynamics are fully endogenous and are described by a partial differential equation akin to the Black-Merton-Scholes formula. We simplify the equation to an ordinary differential equation concerning the total token valuation, subject to an intuitive boundary condition that tokens are worthless for unproductive platforms. We show that the staking ratio, defined as the ratio of aggregate tokens staked in the economy to the total quantity of tokens in supply, proves crucial for token pricing and reward rate determination in equilibrium. Related to the TVL (total value locked) metric widely adopted by practitioners, staking ratio constitutes a new predictor for token price dynamics. UIP is naturally violated because of agents' usage utility which is generally divergent and not reflected in wealth or token price returns, leading to staking-ratio-based trading strategies.

When agents are heterogeneous in wealth and types, the cross-sectional equilibrium reward rate is then determined by the whole distribution of agents, each solving an optimal dynamic control problem taking the market states as given. Under the endogenous evolution of the wealth distribution, the system is captured by coupled stochastic partial differential equations. The staking ratio then involves the weighted average of individual staking choices, which in turn is shaped by the agents' wealth distribution. The resulting reward rate further enters agents' onchain wealth dynamics. Therefore, agents' individual dynamic optimizations interact and co-evolve with the wealth distribution in such an MFG. In the presence of aggregate demand or productivity shocks, the resulting equilibrium is characterized by a so-called master equation, which relates the value function to individual states and also captures influence from the system states with common shocks.

We show that our findings from the baseline model remain robust, and further discuss the impact and evolution of wealth heterogeneity. In particular, we derive analytically the expected stationary wealth distribution, which is Pareto-like. Higher platform productivity provides greater usage benefits for all users while exacerbating their wealth inequality. In a way, our model helps rationalize the documented rising wealth concentration as the Ethereum ecosystem grows (Cong et al., 2022). In general, the tradeoff between aggregate welfare and inequality also depends on staking design, user-type distribution, and

the platform development stage. Next, we numerically examine the impulse responses of wealth distribution and staking ratio to aggregate shocks, e.g., in platform productivity. We observe over-responses in the staking ratio due to the intricate interplay among agents within a dynamically evolving distribution, along with the cascading consequences of the platform's post-shock development.

We derive three main sets of model predictions concerning the economics of staking and its asset pricing implications. First, the staking reward is positively related to the staking ratio in the cross-section. While more staking reduces the reward rate for any given reward quantity, more rewards induce more staking, creating a higher staking ratio for individual agents and in aggregate (which is exacerbated by wealth concentration). Second, expected price appreciation increases with the aggregate staking ratio. Note that both the price drift and staking ratio are functions of platform productivity. In equilibrium, agents stake more when the platform productivity is low, yet that is exactly the time that more wealth can be potentially allocated onto the platform in the future. A higher staking ratio also feeds back to the productivity growth, which increases future demands for tokens and their prices. Third, there are generally predictable excess returns to staking over holding the numéraire, which arise as compensation for the losses in transaction and consumption convenience, in a similar spirit as Valchev (2020). The model thus implies that UIP fails for stakable tokens. We derive the expression for crypto carry and find that higher carry attracts greater staking, generating greater price appreciation and excess return. Because the reward rate mechanically decreases with a greater staking ratio, carry predicts lower excess returns in the time series than in the cross-section.

We test these model implications empirically and find corroborating evidence in a comprehensive dataset covering all major stakble tokens (66 tokens from *StakingRewards* spanning July 2018 to November 2022). We first document that a 10-percentage-points increase in the aggregate reward ratio (e.g., from 10% to 20%) is associated with a 7.79-percentage-point higher staking ratio on average. Moreover, the reward rate has a predictable effect on changes in the staking ratio, with a one percentage point increase in the previous week rate increasing the staking ratio in the following week by about 0.026 percentage points. This property is robust to adding both two-way fixed effect and control variables including market cap and token return volatility. However, its significance decreases with longer time intervals, reflecting to some extent the mechanical downward

adjustment of the reward rate when more tokens are staked (because the same staking rewards have to be divided among more staked tokens).

We next verify that a larger staking ratio indeed predicts greater token price appreciation in subsequent weeks. When the staking ratio increases by one percentage point, the corresponding token price appreciates by 0.066% in the following week. Considering that the variation of staking ratio is often large, especially in the cross-section, this effect is relevant for investment decisions. Crypto market and size factors do not explain the predictive power of staking ratio, which is more related to market liquidity and depth, and reflects the fact that tokens can be commodity-like.[3] Staking reduces the supply of liquid cryptocurrencies, and hence pushes up the token's prices and increases the convenience yields of tokens. This resembles how using commodities as collateral increases the spot price and the convenience yield of the underlying commodities (Tang and Zhu, 2016).

Finally, following the international finance literature, we test if "interest rate" (i.e., reward rate) predicts "currency excess returns" (i.e., token excess return) and find that UIP is indeed violated. We construct a carry trade strategy that goes long high-carry crypto assets and shorts low-carry assets, yielding an annualized Sharpe ratio of 1.60 with weekly rebalancing. Crypto carry predicts excess returns almost one-for-one in the cross-section, with a reduced albeit significant effect in the time series. Intuitively, higher reward rates attract more staking, which persists over the locked period, reducing the reward rates going forward and thus total expected returns, as our model implies.

Literature. Our study adds to the literature on blockchain economics and cryptocurrency markets.[4] In particular, we build on the tokenomics framework of Cong et al. (2021b) and Cong et al. (2021d) to add to emerging studies on Proof-of-Stake protocols (e.g., Fanti et al., 2021; Saleh, 2021; Benhaim et al., 2021) and debates on the environmental and scalability issues associated with Proof-of-Work (PoW) protocols (e.g., Cong et al., 2021a; Hinzen et al., 2019). We also complement the emerging literature on DeFi (e.g., Park, 2021; Cong

[3] Commodities Futures Trading Commission (CFTC) regards cryptocurrencies as commodities, see, e.g., https://www.cftc.gov/sites/default/files/2019-12/oceo_bitcoinbasics0218.pdf.

[4] Existing studies mostly examine issues related to consensus algorithms (Biais et al., 2019; Saleh, 2021), cryptocurrency mining (e.g., Cong et al., 2021a; Lehar and Parlour, 2020), scalability (e.g., Abadi and Brunnermeier, 2018; John et al., 2020), fee designs Easley et al. (2019); Basu et al. (2019); Huberman et al. (2021), DeFi (e.g., Harvey et al., 2021; Capponi and Jia, 2021), ICOs (e.g., Lyandres et al., 2019; Howell et al., 2020), pricing of crypto assets (e.g., Liu et al., 2019; Cong et al., 2021b; Prat et al., 2019), manipulation and regulation (e.g., Griffin and Shams, 2020; Li et al., 2021; Cong et al., 2021c, 2023), or digital currencies (e.g., Gans et al., 2015; Bech and Garratt, 2017; Chiu et al., 2019; Cong and Mayer, 2021).

et al., 2022; Li et al., 2022) by providing a framework for analyzing one of the most prevalent forms of DeFi activity.

The most closely related paper to ours is John et al. (2022) which theoretically examines native PoS crypto assets that serve primarily as investment vehicles, whereas we focus on the platform tokens with a combination of utility flow and investment function while endogenizing agents' dynamic consumption off the network. While both studies demonstrate that the equilibrium staking ratio increases in staking rewards, John et al. (2022) find that staked asset value can exhibit a non-monotonic relationship with block rewards and cause redistribution across agents with divergent trading horizons. We complement this by endogenizing potential adopters' outside option, exploring agents' heterogeneity in wealth and usage preference, as well as considering DeFi staking in addition to the PoS consensus. Also closely related is Jermann (2023) who develops a macrofinance model accounting for both the Ethereum EIP-1559 fee mechanism and its new PoS design, while quantitatively estimating the long-run staking ratio of ETH and the implied money supply. While both studies pin down equilibrium staking considering platform usage value relative to staking benefits, we endogenize the platform productivity process whereas Jermann (2023) endogenizes the token supply. Consequently, our focus is on token pricing instead of monetary policy. Fanti et al. (2021) is the earliest to develop a cash-flow-based valuation framework of PoS cryptocurrencies to understand how the liquidity of validators' holdings, token valuation, and network security relate to one another. Their focus is on long-run transaction fees, we focus on block rewards, endogenous reward rate, and transaction dynamics. Empirically, a recent article by Augustin et al. (2022) characterizes the risk and return trade-offs of yield farming using data from PancakeSwap. We offer likely the first theoretical framework to think about returns to DeFi staking, UIP violations, and crypto carry, with empirical corroborating evidence.

Differing from all aforementioned studies, we employ mean-field game analysis to study agent heterogeneity and its interaction with general forms of staking in steady states and transitions. In particular, we derive onchain and overall wealth distribution in the long run, highlight the trade-off between platform development and equality, and demonstrate how staking facilitates redistribution. Our study not only represents one of the earliest adaptations of mean-field games and the master equation solution in finance, but also innovates by introducing expected steady-states and transition path analysis.

Finally, studies in international finance examine uncovered interest rate parity (e.g., Fama, 1984; Lustig et al., 2019). Carry and its predictability have been analyzed not only for currencies but also for other assets such as equities (e.g., Fama and French, 1998; Griffin et al., 2003; Hou et al., 2011), bonds (e.g., Ilmanen, 1995; Barr and Priestley, 2004), and commodities (e.g., Bailey and Chan, 1993; Casassus and Collin-Dufresne, 2005; Tang and Xiong, 2012). Koijen et al. (2018) applies a general concept of carry and finds that carry predicts returns in both the cross-section and time series. We add by documenting UIP violations and carry premia among cryptocurrencies (and with fiat currencies). Recent empirical studies corroborate our findings by documenting deviations from covered interest parity (Franz and Valentin, 2020) and various forms of carry trades involving crypto derivatives (Christin et al., 2023; Schmeling et al., 2023), and cryptocurrency with loanable programs (Fan et al., 2024). We differ in that our study does not rely on interest rates indirectly inferred from derivative markets or from a particular exchange with lending crypto lending programs. More importantly, we provide a tokenomics theory that can help explain the observations and significant crypto carry profits.

The remainder of this paper is structured as follows. Section 2 sets up a dynamic model of staking and token pricing. Section 3 characterizes a homogeneous-agent equilibrium to illustrate key mechanisms and convey economic intuition. Section 4 solves the full model with agent heterogeneity and derives further insights using the mean-field game analysis. Section 5 introduces the data, stylized facts, and hypotheses to test. Section 6 presents empirical evidence that supports our theory. Section 7 concludes.[5]

2 A Model of Tokenized Economy with Staking

2.1 Model Setup

In a continuous-time economy with infinite horizon, a continuum of agents optimally allocate individual wealth between the offline real economy for consumption and a general digital marketplace (e.g., a tokenized blockchain platform, henceforth referred to as "the platform") where they can conduct peer-to-peer transactions while participating in staking programs for network services and contribution (e.g., consensus recordkeeping, liquidity provision, or improving system security in DeFi protocols).[6] A generic consump-

[5]The online appendices contain the institutional details of staking, introduction to mean-field games (including the derivation of the master equation), numerical procedures, and various extended discussions.

[6]Our findings remain robust under finite horizons, though the numerical procedures have to be modified.

tion good serves as the numéraire and the medium of exchange in the digital network is its native token.

Platform productivity and token price. As in Cong et al. (2021b,d), productivity A_t captures the general usefulness and functionality of the platform, i.e., the convenience users obtain by transacting on the platform using its tokens. We assume that A_t evolve endogenously according to:

$$\mathrm{d}A_t = \mu^A(\Theta_t)A_t\mathrm{d}t, \tag{1}$$

where Θ_t is the endogenous staking ratio, i.e., the ratio of the aggregate number of staked tokens to the total number of tokens, which constitutes a potential state variable that influences token prices and agent decisions. In base layers (pan-PoS consensus protocols) and/or higher layers (DeFi applications and Layer 2 projects), staked tokens contribute to the development of the platform by maintaining node operations, facilitating the achievement of consensus, and increasing the security level of the network, respectively. Therefore, a higher staking ratio typically improves platform productivity, i.e., the drift of A_t is weakly increasing in Θ_t.[7]

Without loss of generality, we denote the token price (in units of the numéraire) as P_t, which is endogenously determined by A_t, the token issuance (total supply) Q_t, as well as external demand shocks (e.g., regulatory changes, market sentiment swings, and noise trading that are independent of (A_t, Q_t)) captured in a Markov stochastic process S_t satisfying $\mathrm{d}S_t/S_t = \mu^S\mathrm{d}t + \sigma^S\mathrm{d}Z_t$ with one source of Brownian innovations $\{Z_t, t \geq 0\}$. In our setting, we treat S_t as a demand shifter so as to understand the impact of aggregate shocks.[8] The token price, $P_t = P(A_t, Q_t, S_t)$, should then follow a general diffusion process with endogenous and potentially time-varying μ_t and σ_t, which we shall solve for:

$$\mathrm{d}P_t = P_t\mu_t\mathrm{d}t + P_t\sigma_t\mathrm{d}Z_t. \tag{2}$$

Agents, adoption, and convenience. We normalize the continuum of (heterogenous)

[7] We could have introduced Brownian shocks to A_t directly, as we did in previous drafts, but they do not add new economic insights and are left out for parsimony and ease of numerical computation. Instead, we examine MIT shocks to A_t in numerical procedures. The effects are similar to aggregate demand shocks. Also, staking could hurt platform productivity if the staker competition is so fierce that fewer stakers participate. Instead of allowing μ^A to potentially decrease in Θ_t, we capture this by explicitly modeling the staking competition.

[8] If one allows S_t to be independently and identically distributed across individuals (e.g., capturing individual sentiments), the problem is further simplified without resorting to the Master Equation, as we show in earlier drafts of the paper. Correlations in indivudal sentiments or preferences can also be used to microfound the aggregate demand shock.

agents to be one unit measure. At time t, their wealth distribution is denoted as $m_t = m(w_t)$, with an absolutely continuous density on $W = [\underline{w}, \infty) \subset \mathbb{R}^+$. Agents indexed by i and characterized by wealth $w_{i,t}$, make consumption-portfolio choices among staking tokens, using tokens on the platform, and holding the numéraire (consumption goods or fiat). A platform user is someone who uses tokens either for staking or transactions.

Users gain convenience from holding tokens and conducting economic activities on the platform. Since staked tokens are locked from the staker's perspective, they can only derive transaction convenience from non-staked (tradable) tokens, which we model similarly as in Cong et al. (2021b,d): for an agent holding x_t (in numéraire, positive) worth of tradable tokens on the platform, she derives a utility flow:[9]

$$\mathrm{d}v(x_t) = \mathrm{d}v_t = x_t^{1-\alpha}(N_t A_t U_t)^{\alpha}\mathrm{d}t - \varphi \mathrm{d}t. \tag{3}$$

With $\alpha \in (0,1)$, the marginal transaction convenience $\frac{\partial v}{\partial x} > 0$ and decreases with x_t. $U_t = U_{i,t} = U(u_{i,t}, w_{i,t}) > 0$ is differentiable in $u_{i,t}$ and reflects transaction needs. $u_{i,t}$ indicates the agent's idiosyncratic transaction preference at time t that is i.i.d. with bounded support. The individual's transaction needs increase with wealth $w_{i,t}$, whereas the type varies the marginal needs. N_t is the measure of staked tokens so that a larger N_t corresponds to an ecosystem with great effectiveness of validators that facilitate onchain transaction. For DeFi staking, especially the ones where token staking has contemporaneous network externality, N_t nicely maps to the concept of Total Value Locked (TVL) and the industry's emphasis on it. In general, we only need N_t to be smoothly increasing in the amount of tokens staked L_t and is zero when no token is staked.

At any time t, agents can choose not to participate, so ad hoc switching costs do not drive our results. Agents adopting the platform need to incur a flow cost φ per unit of time for platform adoption to realize the transaction convenience as the second term in (3) shows. It captures the required effort and attention for participation.[10] Even though we focus on token convenience as a medium of exchange, the reduced-form convenience could also include other utility flows such as governance and voting rights.

Following Bansal and Coleman (1996) and Valchev (2020), the convenience of holding numéraire is reflected in the reduction of transaction costs in consumption. Denote the

[9] In the following, the subscript i is omitted in the general formulation of agents whenever feasible.

[10] It can also include the opportunity cost of earning risk-free interests offline, as in Cong et al. (2021b). If we take this interpretation, then the reward rate needs to be redefined to be in addition to the risk-free rate. For simplicity, we set the risk-free rate to zero.

cost as $\Psi_t = \Psi_t(y_t, n_t, A_t) \geq 0$, where y_t and n_t are consumption and numéraire holdings respectively. Naturally, $\frac{\partial \Psi}{\partial y} \geq 0$, $\frac{\partial \Psi}{\partial n} < 0$. Then $-\frac{\partial \Psi}{\partial n} > 0$ reflects the marginal convenience yield of holding the numéraire. Intuitively, when the platform productivity is lower, the relative convenience of numéraire is higher. Thus, we assume $\frac{\partial \Psi}{\partial A} \leq 0$.

Token transaction convenience and cost enter agents' wealth dynamics rather than utility, for two reasons: first, token convenience flows and transaction costs are typically pecuniary in practice, corresponding to business profits and liquidity costs on real balances, respectively; second, this approach is functionally equivalent to accounting them in the utility function (Feenstra, 1986), and is a standard approach in the literature on convenience yields of bonds (e.g., Valchev, 2020).

Staking and staking rewards. Staking rewards incentivize agents to stake their tokens to either generate consensus records in a base layer or participate in some DeFi program, such as a liquidity pool or insurance pool. In practice, staking rewards come from fees others pay and additional token issuance (emission). To model staking rewards from newly issued blocks, we assume that the total amount of tokens at time t, Q_t, follows a general dynamic process: $\mathrm{d}Q_t = E(L_t, A_t)Q_t\mathrm{d}t$, where E is the "emission rate" function and L_t is the aggregate amount of staked tokens at t.[11] This token supply dynamics reflect not only a potential "inflation" but also the redistribution of onchain wealth between transaction users and token stakers, in a similar spirit to that in John et al. (2022).

We denote the rewards from the transaction fees (e.g., ETH gas) as $F_t = F(L_t, A_t, Q_t)$. F_t depends on the current scale of validators in the ecosystem, as well as the platform productivity that reflects the capacity of processing transactions. Therefore, it also involves L_t and A_t.[12] The total amount of tokens distributed as rewards R_t then becomes:

$$R_t = R(E_tQ_t, F_t) \equiv R(L_t, A_t, Q_t). \tag{4}$$

All staked tokens are fungible and consequently all stakers face an instantaneous reward

[11]The emission schedule is typically public information at the time of staking, and can be at least estimated based on real-time blockchain data (see details in Online Appendix OA1.2). Meanwhile, as Jermann (2023) proves, the equilibrium token supply could be well controlled by the ecosystem design. Therefore, we set the dynamic of Q_t without uncertainty terms. However, E could vary over time.

[12]The gas fee of a transaction in practice entails base and priority fees. The base fee is burned whereas the priority fee is paid to the validator as a tip, which varies across validators and users. L_t, as a proxy of the size of validators, does not equal the number of the validators, but reflects the staking scale competition across the validators, and is usually a good proxy in practice. Online Appendix OA1.2 contains more details.

rate akin to interest rates on bank deposits:[13]

$$r_t \equiv \frac{R(L_t, A_t, Q_t)}{L_t}. \tag{5}$$

Here we only require that R_t being increasing in E and F (since they are both sources of staking rewards) and that r_t being weakly decreases with L_t, as is observed in practice.[14] To capture the cost of node operation and risk of slashing, we assume that stakers incur costs at a rate $c_t < r_t$ proportional to their staking amount.[15]

By Itô's Lemma, someone stakes k_t tokens ($k_t P_t$ dollars) experiences wealth changes:

$$\mathrm{d}(k_t P_t) = k_t \mathrm{d}P_t + P_t(r_t - c_t)k_t \mathrm{d}t = (k_t P_t)[(\mu_t + r_t - c_t)\mathrm{d}t + \sigma_t \mathrm{d}Z_t]. \tag{6}$$

2.2 Agents' Problem and Staking as Optimal Control

Taking as given the system staking reward rate, r_t, each agent with personal wealth w_t decides at time t the consumption rate y_t and a portfolio consisting of l_t numéraire-equivalent amount of staked tokens, x_t numéraire-equivalent amount of tradable tokens and n_t numéraire, where $x_t, l_t, n_t \in [0, w_t]$, $n_t = w_t - x_t - l_t$. The individual stochastic control problem involves using (y_t, x_t, l_t) to optimize a discounted life-time utility:

$$\max_{\{y_s, x_s, l_s\}_{s=t}^{\infty}} \mathrm{E}_t \left[\int_t^{\infty} e^{-\phi(s-t)} \mathcal{U}(y_s) \mathrm{d}s \right]. \tag{7}$$

$\mathcal{U}(y_t)$ is the agent's instant strictly-increasing and concave utility from consumption, ϕ is the discount rate. Because staked tokens cannot be instantaneously liquidated, the agent also faces the budget constraint, $y_t \leq w_t - l_t$. The overall wealth dynamics satisfy:

$$\begin{aligned} \mathrm{d}w_t &= [(x_t + l_t)\mu_t + l_t(r_t - c_t) + v_t - y_t - \Psi_t]\mathrm{d}t + (x_t + l_t)\sigma_t \mathrm{d}Z_t \\ &= f(y_t, x_t, l_t; w_t, r_t, A_t)\mathrm{d}t + g(y_t, x_t, l_t; w_t, r_t, A_t)\mathrm{d}Z_t. \end{aligned} \tag{8}$$

Note that if we allow some of l_t to contribute towards x_t, our framework can be used to study the recent popular "liquid staking" that surpassed decentralized lending in TVL,

[13] For simplicity, we do not model the term structure of staking rewards—the focus of John et al. (2022). In our continuous-time setting, we only need staked tokens to be locked for dt. In our empirical tests, we only require agents to know the next period's reward emission.

[14] For example, many programs fix the total amount of rewards, indicating that staking more tokens reduces the reward per token staked. Generally, L_t decreases F by (i) reducing gas or service fees, since a higher L_t implies the ecosystem has more validators or better capacity of validations, and (ii) generating competition on validating transactions, which depreciates the premium of prioritized transactions (see the definition and operation of priority fees on the official website of Ethereum and blocknative). As for emissions E, L_t weakly decreases E in practice, e.g., the reward rate of Ethereum EIP1559 takes the form of $k/\sqrt{L}$. Otherwise, staking would be more like cooperation and generate quite low liquidity to the ecosystem. Jermann (2023) offers an in-depth analysis on optimal policies across a generalized form of $kL^{-1/s}$.

[15] The time-varying nature of c_t captures the fact that gas fees and risks of slashing could change over time. However, this is not crucial for our key economic insights.

where stakers can get staking rewards without being restricted by the lock-up.[16]

2.3 Staking Ratio, Reward Rate, and Market Clearing

Staking ratio. Define $\theta_{i,t}$ as the individual staking ratio given reward rate r_t at time t,

$$\theta_{i,t} = \theta(r_t, A_t, w_{i,t}, u_{i,t}) = \frac{l(r_t, A_t, w_{i,t}, u_{i,t})}{x(r_t, A_t, w_{i,t}, u_{i,t}) + l(r_t, A_t, w_{i,t}, u_{i,t})} = \frac{l_{i,t}}{q_{i,t}}, \tag{9}$$

where $q_{i,t} = x_{i,t} + l_{i,t}$ is the aggregate individual onchain value. The overall staking ratio, Θ, the ratio of the aggregate number of staked tokens to the total number of tokens, is a resulting control under system states, r, A, and the aggregation of agents' states,

$$\Theta_t = \Theta(r_t, A_t, \{w_{i,t}, u_{i,t}\}_{i\in[0,1]}) = \frac{L_t P_t}{Q_t P_t} = \frac{\int_W l(w_t) m_t(w_t) \mathrm{d}w_t}{\int_W [x(w_t) + l(w_t)] m_t(w_t) \mathrm{d}w_t}. \tag{10}$$

Agents stake taking as given the reward rate, i.e., Θ_t is a continuous founction of r_t. Staking ratio is important because it links individual choices with global states.[17]

Equilibrium reward. According to (5), r_t is influenced by the aggregate stake, L_t, and thus by the aggregation of agents' controls, Θ_t. At the same time, Θ_t is a function of r_t. Then the equilibrium staking reward rate, r_t^*, solves the fixed-point problem:

$$r_t^* = r(\Theta^*(r_t^*, A_t)). \tag{11}$$

Note that the reward rate decreases in Θ_t. We later show that all the agents' staking amount increases with r_t, which then gives a unique equilibrium r_t^*.

Define ρ_t as the staking reward ratio: $\rho_t \equiv \frac{R_t}{Q_t} \Rightarrow r_t = \frac{\rho_t}{\Theta_t}$. Since ρ_t has a one-to-one correspondence to the equilibrium r_t^*, the equilibrium staking ratio can also be represented as $\Theta^*(\rho_t, A_t)$. In our subsequent comparative statics analysis, we use $\Theta^*(\rho_t, A_t)$, while retaining the notation $\Theta(r_t, A_t)$ to highlight the output of agents' choices. When the system is at the equilibrium, $\Theta^*(r_t^*, A_t) = \Theta^*(\rho_t, A_t)$.

Token market clearing. The total quantity of tokens Q_t is equal to the sum of token holdings affected by the noisy demand shock S_t:[18]

$$Q_t P_t = \int_W (x_t + l_t) m_t \mathrm{d}w_t S_t. \tag{12}$$

[16] https://www.coindesk.com/markets/2023/02/27/liquid-staking-replaces-defi-lending-as-second-largest-crypto-sector/. This is achieved through wrapped tokens that allow unsupported assets to be traded, lent, and borrowed on DeFi platforms. Our earlier example of Solana is supported by liquid staking platforms such as Lido, as of 2022.

[17] For notational simplicity, $\{w_{i,t}, u_{i,t}\}_{i\in[0,1]}$ are omitted in the remainder of the paper.

[18] Note that from (10) and (12), $P_t L_t = P_t Q_t \Theta_t = \int_W l_t m_t \mathrm{d}w_t S_t$. The token price simultaneously clears both the staking and non-staking markets to rule out arbitrage.

3 Equilibrium Characterization and Implications

To convey core economic intuitions, we first solve for a baseline Markovian equilibrium with homogeneous agents each having the same type u and non-negative (but finite) initial wealth, and set the transaction threshold to be zero. In a symmetric equilibrium, the evolution of individual wealth is the same across agents, which is denoted by w_t. Section 4 derives the equilibrium under the general settings with agent heterogeneity.

For homogeneous agents, we leave out the distribution m and write the indirect utility function (given that Q_t does not enter, as we show later):

$$J(t, w_t, A_t, r_t) = \max_{\{y_s, x_s, l_s\}_{s=t}^{\infty}} \mathrm{E}_t \left[\int_t^{\infty} e^{-\phi(s-t)} \mathcal{U}(y_s) \mathrm{d}s \right]. \tag{13}$$

We then derive the Hamilton-Jacobi-Bellman (HJB) equation:

$$\phi J(t, w, A, r) = \max_{\{y,x,l\}} \left\{ \mathcal{U}(y) + f(y, x, l; w, r, A) \frac{\partial J}{\partial w} + \frac{1}{2} g(y, x, l; w, r, A)^2 \frac{\partial^2 J}{\partial w^2} + \mu^A(\Theta) A \frac{\partial J}{\partial A} \right\}, \tag{14}$$

where we leave out the time subscript, t, for simplicity.

3.1 Rewards and Staking Activities

We start by analyzing the agents' optimal decisions. At the instant of decision-making, the agent takes the reward rate r_t as given. Intuitively, the marginal transaction convenience is decreasing with x_t, the value of tradable token held.

Proposition 1. ***Optimal individual staking.*** *For an agent with wealth w_t and type u under system states $\{A_t, r_t\}$, the optimal individual staking ratio, θ_t^*, is unique and satisfies:*

$$\theta_t^* = \max \left\{ 0,\ 1 - \frac{N_t A_t U_t}{q_t} \left(\frac{1-\alpha}{r_t - c_t} \right)^{\frac{1}{\alpha}} \right\}. \tag{15}$$

The reward rate, r_t, appears in the denominator, implying that the individual staking ratio is weakly increasing in r_t. When the marginal transaction convenience becomes smaller than the staking reward rate, i.e., $r_t - c_t > (1-\alpha)(\frac{N_t A_t U_t}{x_t^*})^{\alpha}$, the agent starts to stake excess tokens. Because for an individual, the benefits of staking are constant, an interior solution exists. Substituting the agents' individual optimal choices into (10), we obtain the aggregate staking ratio $\Theta(r_t, A_t, w_t, u)$. Market clearing guarantees the existence of the equilibrium staking ratio, which is weakly increases with reward rate r_t. $\{r_t, \Theta(r_t, A_t, w_t, u)\}$ satisfies (10) and (11). The equilibrium is determined by the aggregate staking reward ratio ρ_t. The following proposition states that higher aggregate staking rewards attract agents to stake more, leading to higher staking ratios.

Proposition 2. ***Equilibrium staking ratio.*** *The optimal onchain wealth, $q_t = x_t + l_t$, is unique and positive. The equilibrium overall staking ratio, Θ_t^*, is unique and satisfies:*

$$\rho(\Theta_t^*) = \Theta_t^* \left\{ (1-\alpha) \left[\frac{N(\Theta_t^*) A_t U_t}{q_t (1-\Theta_t^*)} \right]^{\alpha} + c_t \right\}. \tag{16}$$

In particular, when the reward ratio ρ is treated as a system state, higher ρ leads to higher aggregate staking ratio in equilibrium, i.e., $\forall \rho' > \rho > 0$,

$$\Theta^*(\cdot, \rho') > \Theta^*(\cdot, \rho). \tag{17}$$

Proposition 2 gives a general characterization of how aggregate staking reward affects staking ratio in equilibrium. The result (16) applies to both cross-sectional comparison and time series analysis. Its right-hand side reveals two ways wealth inflows affect transaction conveniences. The inflow from expected price appreciation has a diminishing marginal transaction convenience on the one hand, and thus moves to the stake pool, increasing the network effect therefore increasing the convenience benefits, ultimately making the two forces in balance. Importantly, (17) generates comparative statics that the staking ratio is increasing in the staking reward ratio.

3.2 Staking Ratio & Price Dynamics

We link staking activities to token prices. In general, token price appreciates when more agents' wealth flows into the platform due to high productivity and thus large transaction convenience, or greater staking. We are interested in the drift term μ_t of token prices, which depends on both the platform states and the agents' control cross-sectionally. For simplicity, we omit the subscript t when there is no ambiguity.

By Proposition 2, a participating agent's optimal onchain wealth, q_t, is unique. Combining with the market clearing condition, (12), we obtain:

$$0 = \left(\mu + r(\Theta) - c + \frac{\partial \Psi}{\partial n} \right) \frac{\partial J}{\partial w} + \frac{PQ}{S} \sigma^2 \frac{\partial^2 J}{\partial w^2}, \tag{18}$$

where $\frac{\partial \Psi}{\partial n}$ may also be a function of P_t and Q_t. Further, as P_t is endogenously given by A_t, Q_t, and S_t, μ_t and σ_t are also endogenously determined. Applying Itô's Lemma and matching the coefficients to (2), we obtain:

$$\mu_t = \frac{1}{P_t} \left(\frac{\partial P_t}{\partial A_t} A_t \mu_t^A + \frac{\partial P_t}{\partial Q_t} Q_t E_t + \frac{\partial P_t}{\partial S_t} S_t \mu_t^S + \frac{1}{2} \frac{\partial^2 P_t}{\partial S_t^2} (S_t \sigma_t^S)^2 \right), \quad \sigma_t = \frac{1}{P_t} \frac{\partial P_t}{\partial S_t} S_t \sigma_t^S. \tag{19}$$

Note that $\frac{\partial P_t}{\partial Q_t}$ is negative, thus the emission adds negatively to price drifts. In other words, emissions of staking rewards cause "inflation," a concern token holders well recognize.

Combined with (19) and the previous propositions, (18) can be rearranged as a partial differential equation (PDE) with respect to A_t and Q_t, i.e., with $I = -\frac{\partial_{ww} J}{\partial_w J}$,

$$\begin{aligned} 0 = & \frac{\partial P}{\partial Q} EQ + \frac{\partial P}{\partial A} A\mu^A(\Theta^*) \\ & + \left[\frac{\partial P_t}{\partial S_t} S_t \mu_t^S + \frac{1}{2} \frac{\partial^2 P_t}{\partial S_t^2} (S_t \sigma_t^S)^2 - \frac{QI}{S} \left(\frac{\partial P}{\partial S} S_t \sigma^S \right)^2 \right] + \left[r(\Theta^*) - c + \partial_n \Psi \right] P, \end{aligned} \tag{20}$$

(20) differs from a Black-Scholes-type PDE. First, the "theta" term in Black-Scholes (BS) equation reflecting the variation of the derivative value over time is absent in (20). Instead, the term $\frac{\partial P}{\partial Q} EQ$ captures the expected inflation from token issuance. Second, since A_t, the productivity that drives token price, is not tradable, the coefficient of $\frac{\partial P}{\partial A}$ is $A\mu^A$ instead of zero.[19] Moreover, there is a "flow" term, $\left[r(\Theta^*) - c + \partial_n \Psi \right] P$, that reflects the excess gain from staking rewards offsetting the staking cost and convenience loss. $\partial_n \Psi$, indicating the loss of numéraire convenience, is typically negative and could be a function of A_t. $\left[\frac{\partial P_t}{\partial S_t} S_t \mu_t^S + \frac{1}{2} \frac{\partial^2 P_t}{\partial S_t^2} (S_t \sigma_t^S)^2 - \frac{QI}{S} \left(\frac{\partial P}{\partial S} S_t \sigma^S \right)^2 \right]$ captures the impact of external risk, where the later term features the impact of agents' risk aversion, and does not appear in the risk-neutral BS equation.

(20) is a PDE involving derivatives w.r.t. A_t, Q_t, and S_t, which is hard to solve. Recall the market clearing condition (12) can be alternatively represented as a condition on aggregate wealth allocated to the platform, which relates to the emission rate instead of Q_t. Substituting the separable form of P_t into (20) and pining down the relevant functions, we convert the PDE into an ordinary differential equation (ODE).

Proposition 3. ***Token price and dynamic.*** *P_t is separable, $P_t = \frac{V(A_t) S_t}{Q_t}$, where $V(A_t)$ captures the aggregate onchain wealth allocated. $V(A_t)$ satisfies the following ODE given log agent utility, Ψ_t linear in numéraire holding, and zero sentiment (noise) drift μ^S,*[20]

$$0 = V'(A_t) A_t \mu^A(\Theta_t) + \left[r(\Theta_t) - c_t - E(\Theta_t, A_t) + \Psi_n(A_t) \right] V(A_t) - \frac{[\sigma^S(A_t) V(A_t)]^2}{w_t}, \tag{21}$$

where $\Theta_t = \Theta(A_t, V(A_t))$ and satisfies (16), w_t indicates the aggregation of agents' wealth. Given the boundary condition $\lim_{A_t \to 0} V(A_t) = 0$, the ODE has a unique solution.

The economic implications of (21) are similar to our previous discussion of (20). Under a fixed inflation rate, in equilibrium, the expected price drift, μ_t, and staking ratio, Θ_t, are

[19] If the fundamental productivity were tradable, the coefficient of $\frac{\partial P}{\partial A}$ would have been $r^f A$, where r^f is the risk-free rate and normalized to zero.

[20] We specify a log utility given that the level risk aversion is not our focus, despite the numerical simulations verify that the model implications hold under general CRRA utilities.

both functions of platform productivity, A_t. The boundary condition represents that a useless platform attracts no adopters.

Solving the ODE, we find that the equilibrium staking ratio Θ_t and expected price appreciation μ_t are positively related (Figure 1). Because staking ratio can be computed from onchain information, it can help predict price changes.

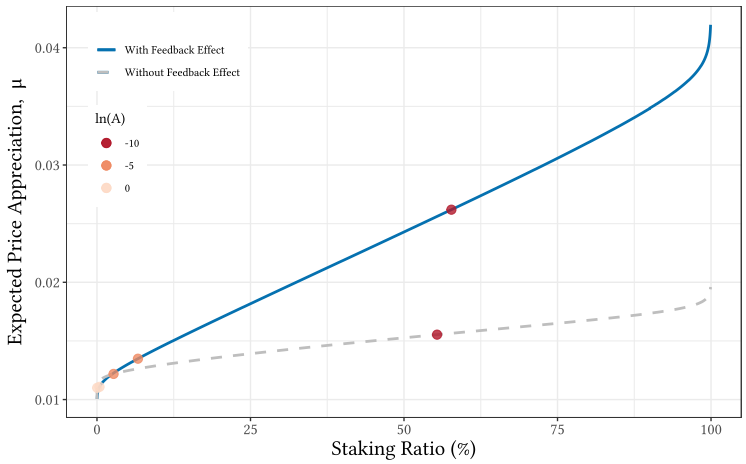


Figure 1: Staking ratio and price dynamics.
This graph shows the joint relationship between the system staking ratio Θ_t and the price drift μ_t. The blue curve is the case where the staking ratio feeds back the platform productivity A_t process (the main model), while the gray curve shows the case for comparison where the feedback effect does not exist. The colored scatter points mark the corresponding points (Θ_t, μ_t) for different values of $\ln(A_t)$ respectively.

There are two main economic driving forces. The first force comes directly from productivity A_t. On the one hand, μ_t decreases in A_t. As A_t grows, agents allocate more wealth on the platform and less off-chain wealth, thus the potential future price appreciation is reduced, which generates the similar user-base stabilizing effect of tokens as Cong et al. (2021b). On the other hand, Θ_t also declines in A_t, because higher A_t results in a larger transaction convenience. Therefore, the joint dynamics of μ_t and Θ_t exhibit a positive relationship. This mechanism also explains the steep slope of the curve when the staking ratio is low. Such force does not rely on the feedback mechanism but in general a market phenomenon. Therefore, we can see from Figure 1 that even without the feedback effect (the grey curve), μ and Θ are positively related.

The second is the feedback effect of staking on the A_t process. As (1) and (19) show, a high staking ratio increases the productivity drift μ_t^A, and then leads to a large price drift μ_t. This force accounts for the role that staking plays in platform growth. In PoS, the system state with a relatively high staking ratio implies a strong network of highly

engaged validators, so that the consensus and confirmation are efficiently reached. As for high-layer staking economy such as DeFi applications, with a certain capital value, a high staking ratio relates to a high TVL (total value locked), which is recognized as improving the security level of the platform. For both layers of the staking economy, the staking ratio positively impacts the growth of platform productivity A_t through the above-mentioned paths respectively, therefore resulting in a greater drift. As a reflection of the value of the platform, the price drift increases accordingly to (19). Online Appendix OA3.5 discusses further the implications of the feedback effect.

3.3 Token Excess Returns, UIP Violation, and Crypto Carry

Each token holder takes on the risk of token price fluctuation and loses the convenience of holding the numéraire. Denote the expected financial excess return of unit staked token over the numéraire by λ_t, $\lambda_t \equiv \mathrm{E}_t\left[\mathrm{d}P_t + P_t r^{staked\ token}\right]/P_t = \mu_t + r_t - c_t$.

Proposition 4. ***Predictable excess return.*** *The unit excess return, λ_t, satisfies*

$$\lambda_t = -\frac{\partial \Psi}{\partial n_t} + q_t^* \sigma_t^2 I > 0, \tag{22}$$

where $-\frac{\partial \Psi}{\partial n_t} > 0$ is the marginal convenience of holding numéraire, q_t^* the optimal onchain wealth, and $I = -\frac{\partial_{ww} J}{\partial_w J} > 0$ the risk aversion. Excess returns are predictable by a positive λ_t.

Proposition 4 follows directly from the agents' optimization. The positive RHS of (22) implies predictable excess returns that arise as compensation for convenience losses and volatility risk. This phenomenon closely relates to the uncovered interest rate parity (UIP) in the foreign exchange market, since the token price in numéraire corresponds to the exchange rate, whereas the staking reward rate corresponds to the interest rate. UIP implies that the expected returns on default-free deposits across currencies are equalized, and thus the expected excess return λ_t should be zero.[21] However, (22) violates the UIP owing to the presence of two positive terms. First, when the relative convenience of numéraire increases, staked token is compensated with a higher financial return. This interpretation of UIP violation shares similar ideas with Valchev (2020); Jiang et al. (2021)'s

[21] In relevant research on currency UIP, the more common notation reads $\lambda_t = \mathrm{E}_t\left[\mathrm{d}S_t/S_t + i_t^{foreign} - i_t^{local}\right]$, where S_t is the log exchange rate (foreign currency units per unit of local currency). The corresponding terms of S_t, $i^{foreign}$ and i^{local} are P_t, $r_t - c_t$ and the numéraire risk-free rate r_t^f (normalized to zero), respectively. In related empirical works, r_t^f is not always zero, and the time is discrete. The corresponding equation of UIP reads $\mathrm{E}_t\left[\log P_{t+1} - \log P_t\right] = r_t^f - (r_t - c_t)$.

explanation of the UIP puzzle in classical asset types such as bonds. The second term on the R.H.S represents the impact of volatility risk.[22]

In a nutshell, the excess return of holding the token comes from both the staking reward and the price appreciation of tokens. But there is no free lunch. The excess return reflects the compensation for the missed convenience of holding the numéraire for consumption. This observation is general because convenience is a relative concept between any two assets. What this implies is that based on the same numéraire, the expected excess returns can be different for different tokens. Moreover, by using any of cryptocurrencies as numéraire, we can verify that UIP fails in general in the cryptocurrency market.

UIP violations naturally lead to profitable currency carry trades, which go long in baskets of currencies with high interest rates and short in low ones. In fact, carry is a general concept that applies in a host of asset classes, e.g., equities, bonds, commodities, Treasuries, credits, and options (Koijen et al., 2018). Its predictability of excess returns and investment performance are widely documented and studied (e.g., Lustig et al., 2014; Bakshi and Panayotov, 2013; Burnside et al., 2011; Menkhoff et al., 2012; Koijen et al., 2018; Daniel et al., 2017). As a direct application to the violation of UIP, our model implies profitable crypto carry trades, where following Koijen et al. (2018), crypto carry is similarly defined as currency carry:

$$\text{carry}_t \equiv \frac{r_t - c_t - r^f}{1 + r^f}. \tag{23}$$

4 Agent Heterogeneity and Mean-Field Equilibrium

We now solve the full model under agent heterogeneity in both wealth and usage need/preference for the platform. We find that all aforementioned model implications hold. We also derive the expected stationary wealth distribution, which exhibits a Pareto-like shape. While higher platform productivity offers greater transaction convenience, inequality also increases. We further subject the economy to aggregate shocks to examine its response dynamics, thereby elucidating the competitive dynamics among agents, rationalizing the direct and indirect impact paths, etc.

[22] This explanation is related to studies using term structure models (e.g., Bansal, 1997; Lustig et al., 2019), where the difference between domestic and foreign bond risk premia, expressed in domestic currency, is determined by the volatility difference of the permanent components of the stochastic discount factors. In addition, alternative relevant explanations for the UIP violation have been proposed in previous studies, ranging from time-varying risks including liquidity and volatility (e.g., Bekaert, 1996; Verdelhan, 2010; Lustig et al., 2011; Gabaix and Maggiori, 2015) to peso problems (e.g., Burnside et al., 2011).

Intuitively, poor agents with greater usage preference of the platform may stake less, which in turn affects the aggregate staking ratio and token pricing. Further, introducing wealth heterogeneity into the dynamic process enables us to discuss the evolution of the wealth distribution and its long-run outcomes, and thus inequality and redistribution.

4.1 The MFG Approach and Derivation of the Master Equation

Wealth heterogeneity and the mean-field game. With wealth heterogeneity, the density m_t enters the determination of staking ratio and reward rate as (10) and (11) show. Since agents solve their optimal control based on staking reward rates, m_t enters the value function as an additional argument. For such a system involving the evolution of distribution, we employ the so-called "backward-forward MFG system" for the system with a backward HJB and a forward FP equation (law of motion).[23] Because the external aggregate (common) shock hits every agent, it adds randomness to the distribution so that $\{m_t\}_{t\geq 0}$ is a random flow of measures.[24] The resulting stochastic PDE system has been thoroughly studied on its mathematical properties and applications, especially the equivalent system defined by a so-called "master equation." In the following, we build upon Cardaliaguet et al. (2019) and Bilal (2023) to characterize the system with a master equation. For notational simplicity, we omit unused arguments of a function whenever feasible.

The standard time-dependent formulation of the coupled stochastic PDEs reads:

$$
\begin{cases}
\phi J_t \mathrm{d}t - \mathrm{d}J_t = \left\{ \max\limits_{y,x,l} \left[\mathcal{U}(y;w,m) + f(x,l;w,m)\dfrac{\partial J_t}{\partial w} + \dfrac{1}{2}g^2(x,l;w,m)\dfrac{\partial^2 J_t}{\partial w^2} \right] \right. \\
\qquad\qquad \left. + \mu^A A \dfrac{\partial J_t}{\partial A} - \dfrac{\partial g(w,m)v_t}{\partial w} \right\} \mathrm{d}t - v_t \mathrm{d}Z_t \\
\qquad\qquad \equiv \left\{ \mathcal{U}(w,m) + L(w,m)[J] + \mu^A A \dfrac{\partial J_t}{\partial A} - \dfrac{\partial g(w,m)v_t}{\partial w} \right\} \mathrm{d}t - v_t \mathrm{d}Z_t, \\
\mathrm{d}m_t = \left[-\dfrac{\partial}{\partial w}\left(f(w,m)m_t(w)\right) + \dfrac{\partial^2}{\partial w^2}\left(\dfrac{1}{2}g^2(w,m)m_t(w)\right) \right] \mathrm{d}t - \dfrac{\partial}{\partial w}\left[g(w,m)m_t(w)\mathrm{d}Z_t\right] \\
\qquad\qquad \equiv L^*(w)[m]\mathrm{d}t - \dfrac{\partial}{\partial w}\left[g(w,m)m_t(w)\mathrm{d}Z_t\right],
\end{cases}
\tag{24}
$$

[23] The mean field game (MFG) theoretical approach has been developed by Lasry and Lions (2007); Cardaliaguet et al. (2019), and applied in macroeconomics (e.g., Achdou et al., 2022) and related Bitcoin mining competition (Bertucci et al., 2020). Online Appendix OA3.1 provides a detailed review.

[24] Let us compare aggregate shocks with another class of shocks that are also usually discussed, where each agent has an independent individual uncertainty. Although every agent randomly moves from her initial state, the stochasticity of directions and displacements is smoothed out after the summation of numerous agents, resulting in a deterministic evolution of the distribution. In contrast, the aggregate shock here randomly shifts system states and thus affects all agents systematically, despite the different exposures.

where the second equal signs of the two equations define the notations for the optimized situation, including the utility $\mathcal{U}$, wealth drift f and diffusion g. $L(w)[J]$ is the Dynkin operator w.r.t. the individual state w. $L^*(w)[m]$ is also a useful operator, in which the star script reminds it is the adjoint operator of L. v_t is a function for penalty.[25]

Regarding this system, the key of the master equation approach is to notice that the time-dependence of the value function $\mathrm{d}J_t$ itself in (24) comes totally from the time-dependence of m_t.[26] As shown below, it treats J_t as $J(\cdot, m_t)$ and replaces $\mathrm{d}J_t$ by the FP equation and the infinite-dimensional Itô's Lemma. As Bilal (2023) points out, though additional definitions introduced, the master equation brings substantial ease for discussion on economic insights, since it captures the system to be a Markovian representation of the agents' problem by a single equation. In this economy, the master equation reads:

$$\begin{aligned}\phi J(w,m,A) = {} & \mathcal{U}(w,m) + L(w)[J] + \mu^A A \frac{\partial J}{\partial A} + \frac{\partial}{\partial w}\left[g(w)\int \frac{\delta J}{\delta m}(w,w',m)\frac{\partial g(w')m(w')}{\partial w'}\mathrm{d}w'\right] \\ & + \int \frac{\delta J}{\delta m}(w,w',m)L^*(w')[m]\mathrm{d}w' \\ & + \frac{1}{2}\int\int \frac{\delta^2 J}{\delta m^2}(w,w',w'',m)\frac{\partial g(w')m(w')}{\partial w'}\frac{\partial g(w'')m(w'')}{\partial w''}\mathrm{d}w'\mathrm{d}w''.\end{aligned} \tag{25}$$

Without heterogeneous wealth and aggregate shocks, the master equation degenerates into the HJB equation, which corresponds to the first three terms on the right-hand side. The next three terms capture the decomposition of the impact of m_t on the forward evolution of the MFG system: the first is derived from the penalty, i.e., it captures how the agents evaluate the impact of the realization of the evolutionary uncertainty of the wealth distribution. The second represents the impact of the "deterministic" changes in the wealth distribution. Similar to the volatility risk (acts on wealth) of classical uncertainties, the evolutionary uncertainty of the distribution is volatile and results in a risk (the second-order term w.r.t. m), as the last term in (25) reflects.[27] The derivation is shown in

[25] With the randomness in forward-backward SDEs, the HJB equation needs to involve a martingale as penalties to guarantee the solution is indeed adapted (according to the theory regarding forward-backward SDEs, see e.g., Pardoux and Râşcanu, 2014).

[26] Put differently, $J_t = J(w_t, A_t, m_t)$ and the total derivative would include the time-dependence of all the state arguments. While in the HJB, the time-dependence of w_t and A_t have already been accounted for.

[27] Note that extra arguments w', w'' enter the derivatives. They are generated from infinite-dimensional Itô' Lemma: $\mathrm{d}J_t = \langle \frac{\delta J}{\delta m}, \mathrm{d}m_t\rangle + \frac{1}{2}\langle \mathrm{d}m_t | \frac{\delta^2 J}{\delta m^2} | \mathrm{d}m_t\rangle$. The inner product is defined in the appropriate functional space, $\langle f(x), g(x)\rangle = \int f(x)g(x)\mathrm{d}x$. $\langle h | f(x,x') | g\rangle = \langle\langle f(x,x'), h(x)\rangle, g(x')\rangle$. The derivatives w.r.t. m refer to the Fréchet derivatives applied in infinite dimension, and appear to have extra arguments, i.e., $\frac{\delta J}{\delta m} = \frac{\delta J}{\delta m}(w,w',m)$, $\frac{\delta^2 J}{\delta m^2} = \frac{\delta^2 J}{\delta m^2}(w,w',w'',m)$. The rigorous definitions are described in Online Appendix OA3.1.

Online Appendix OA3.2. The extra terms imply that aggregate shocks not only affect any agent i directly, also affect other agents, whose responses in turn affect agent i indirectly.

4.2 Implications of Wealth Heterogeneity

Individual's optimal staking. Heterogeneous agents differ in the trade-off between staking and convenience gains they face, some even quit the staking pool.

Proposition 5. ***Individual staking under heterogeneity.*** *For agent i with wealth w_t and usage preference $u_{i,t}$, the optimal q_t^* is unique. The optimal individual staking ratio θ_t^* is heterogeneous w.r.t. agents' transaction needs $U_t = U(u_t, w_t)$ and satisfies:*

$$\theta_t^* = \max\left\{0, 1 - \left(\frac{1-\alpha}{r_t - c_t}\right)^{\frac{1}{\alpha}} \frac{N_t A_t U_t}{q_t^*} \mathbb{I}\{U_t > U_0\}\right\}, \tag{26}$$

where $\mathbb{I}$ is an indicator function, $U_0 = \frac{\varphi}{N_t A_t}\left(\frac{r_t - c_t}{1-\alpha}\right)^{\frac{1-\alpha}{\alpha}}$.

The indicator function captures the impact of the threshold φ. The unchanged, however, is all the agents weakly increase their staking under a higher reward rate. Figure 2 visualizes this implication and the determination of the cross-sectional equilibrium. As the aggregation of all the agents, the overall staking ratio (the blue curve) weakly increases with the given reward rate. Further, the resulting staking ratio leads to a new system reward rate (the gray curve). Then the equilibrium falls at the intersection.[28] In addition, we can see that big whales enter the staking pool first as the reward rate increases, since their marginal transaction needs are relatively lower. From another perspective, the retail agents are less likely to be interested in staking, since the staking reward, also their contribution to the platform by staking are relatively negligible.

System equilibrium states and redistribution. The qualitative implications of the model remain intact despite the presence of heterogeneous agents.[29] However, the crucial difference is that the wealth distribution plays a pivotal role as an additional argument in shaping all the cross-sectional equilibrium states, leading to redistribution. Specifically, heterogeneity in wealth causes agents to have different tradeoffs between transaction usage and staking rewards, leading to different individual staking ratios. Agents

One can roughly compare them with Jacobian and Hessian in the finite-dimensional case respectively.

[28] When it falls in the gray area, some agents may not stake and only transact on the platform.

[29] The essence that generates this assertion is that the system states are still calculated as aggregations across the crowd, whereas agents have the same direction of shifts in staking relative to different reward rates as mentioned above. Online Appendix OA3.3 provides a detailed discussion on the derivation process.

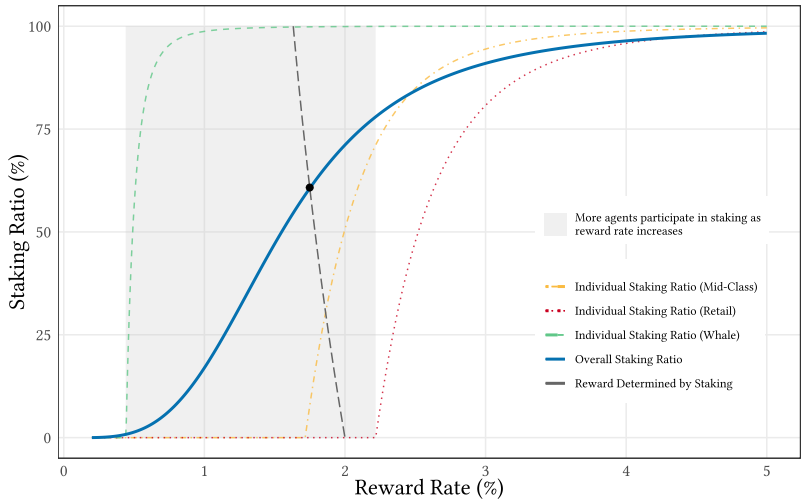


Figure 2: Individual staking decisions and equilibrium staking ratio.
The three dashed curves plot the individual staking ratios of retails (who own little wealth), rich whales, and the middle-class. The blue curve is the aggregated overall staking ratio. The staking ratio further leads to the resulting reward rate as the gray curve shows. Then, the equilibrium is the solution to the fixed point problem as visualized by the intersection of the blue and gray curves.

who value usage benefits more or have less wealth tend to stake less, and consequently suffer from emission inflation, whereas stakers, especially the wealthy ones, accumulate more tokens, and thus onchain wealth. This insight shares the same spirit as in John et al. (2022), though the agent heterogeneity modeled there is different, and the redistribution is between short-term and long-term holders. With log utility, onchain redistribution also corresponds to an overall wealth redistribution in equilibrium.

4.3 Expected Stationary State

The cross-sectional equilibrium exists under any given wealth distribution, while we further consider the steady state from the time dimension. With the continuous arrival of aggregate shocks, we should consider the expected steady state, $\mathbb{E}(\mathrm{d}m_t) = 0$. It reflects the evolution on average and directly corresponds to the deterministic onchain steady state if there is no realized external shock. As discussed in Appendix OA3.5, if the agents obtain log utilities, $U_{i,t} = U(u_{i,t}, w_{i,t}) = u_{i,t} w_{i,t}$, and the random user type $u_{i,t}$ is independent of wealth, then all the agents' wealth change linearly. It is widely known that the resulting stationary distribution is a Pareto distribution.[30] We relax the above requirement

[30] When the agent's wealth changes linearly over time, the law of motion generates a Pareto distribution (e.g., Wold and Whittle, 1957; Gabaix, 2009). It is worth pointing out that we make a methodological contribution by avoiding the common assumption in the literature of a representative agent that gives a deterministic evolution of the distributions.

by allowing for correlation between u and w, e.g., $u_{i,t} \sim \bar{u}\mathcal{N}(w_t^{\beta-1}, \sigma_X^2)$, where $\bar{u}$ captures the tendency to transact (as opposed to stake). Then from the perspective of the whole crowd, there are different marginal benefits between staking (still constant return to scale given the reward rate r) and transaction (expectations affected by wealth). The following proposition solves for the general expected stationary distribution and shows it depends on the platform productivity.

Proposition 6. ***Expected stationary distribution.*** *The agents' wealth distribution in the expected stationary state is:*

$$m(w) = c_0 \frac{1}{w^{\phi\mathcal{V}}} \exp\left(\mathcal{V}\bar{u}\alpha\left(\frac{1-\alpha}{r-c}\right)^{\frac{1-\alpha}{\alpha}} NA\frac{w^{\beta-1}}{\beta-1}\right), \tag{27}$$

where $\mathcal{V} = \frac{2\sigma^2}{(\mu+r-c-\Psi_n)^2}$, $(\mu, r, \Psi_n, \sigma, N)$ *are all functions of* A *at the cross-sectional equilibrium.* c_0 *is a normalization constant such that* $\int_{\underline{w}}^{\infty} m(w)\mathrm{d}w = 1$. *As* $\beta \to 1$, *the distribution becomes a Pareto distribution with tail parameter* $a = \left[\phi - \alpha\left(\frac{1-\alpha}{r-c}\right)^{\frac{1-\alpha}{\alpha}} NA\bar{u}\right]\mathcal{V} - 1$.

We further investigate how the platform state influences the stationary wealth distribution as Figure 3 shows.[31] The distributions are Pareto-like, with a large amount of wealth concentrated in a few people. Panel A tells us that as the platform productivity increases, the distribution involves a fatter tail, implying the rising wealth inequality. Higher platform productivity generates larger benefits than the off-chain activities, while in this process, the richer has larger onchain allocation, and gains more wealth growth. Meanwhile, higher productivity increases the transaction convenience of all agents, which explains why many retail agents still use the platform.

Similarly, as demonstrated in Panel B, higher tendencies to stake (lower $\bar{u}$) reduce inequality. In fact, both user type ($\bar{u}$ and β) and staking design (e.g., ρ) can alter the level of inequality. Platforms without staking or with less generous staking rewards (lower ρ) would not have as much staking participation, and therefore exhibit higher inequality.

Wealth inequality and total welfare. Just like in the real economy, inequality can be a double-edged sword. On the one hand, it results in a lower equilibrium staking reward rate and squeezes out retail agents from staking, which reduces their welfare. On the other hand, inequality is correlated with high platform productivity (and thus greater average

[31] To make the situations comparable, we normalize the wealth to the multiplier of the lowest onchain wealth in corresponding cases. Since in different cases, the agents have co-movements in on/off-chain trade-offs. The normalized onchain wealth allows us to focus on inequality issues.

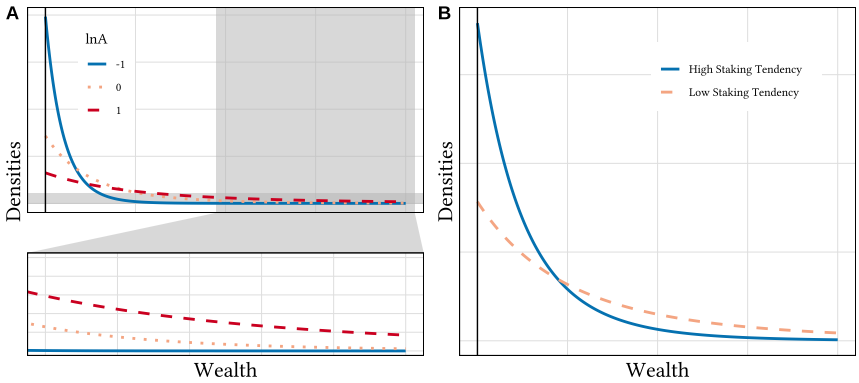


Figure 3: Expected stationary wealth distribution.
This graph shows the stationary wealth distribution under different platform states. The wealth is normalized to be the multiplier of the lowest wealth in each case. Panel A shows the comparison among different levels of platform productivity. Panel B compares different overall staking tendencies, which is examined by adjusting $\bar{u}$, the general scale of transaction demand in user types.

user utility). Agents may still want to join a well-developed platform for usage, even in an economy with large onchain and overall wealth inequality.

These implications further help us understand the different stages of platform growth. In the initial stage, the tokenized economy is often accompanied by relatively low inequality. Participating in staking can attract excess capital into the network and promote its development. As the platform grows, a few rich agents with larger marginal benefits in staking than transacting dominate staking activities, whereas other agents hold tokens mostly for onchain usage and speculation. This helps to rationalize why in practice, once platforms become large, a division of labor among participants emerges where validation is often done by prominent asset owners such as exchanges, who provide consensus through PoS while earning staking revenues.

4.4 System Trajectories and Aggregate Shocks

According to the master equation (25), agents do not directly respond to any specific shocks, but "internalize" them through the penalty term as well as the volatility risk of penalty. To numerically solve (25), we obtain the value functions under shock series z, denoted as $J(w, m, A, z)$, and simulate all the possible series under the given stochastic process to derive the origin $J(w, m, A)$. This idea offers additional potential to examine the evolutionary paths under certain realized series of shocks. To solve J, we follow the perturbation approach introduced in Bilal (2023), as detailed in Online Appendix OA2.

Figure 4 shows the impulse response paths when there are different series of shocks.

Panel A simulates a booming market with positive external demand shocks.[32] The direct and initial effect is a universal wealth growth across all agents, leading to a decrease in the share of retail agents, and consequently an increase in the overall staking ratio. As time goes on, the growing productivity generates more benefits from transaction and reduces the staking ratio. Panel B simulates a declining market with negative demand shocks, in which the response paths are nearly symmetric. Panels C and D reveal the periods with low and high oscillations. The whole system has immediate and violent response to each round of external fluctuations. This is because the latest arriving shock brings the direct impact, while the previous opposite shock only affects the current state indirectly by the changes in the distribution.[33] Further, the changing nature of productivity obtains quantitative-level differences among the rounds.

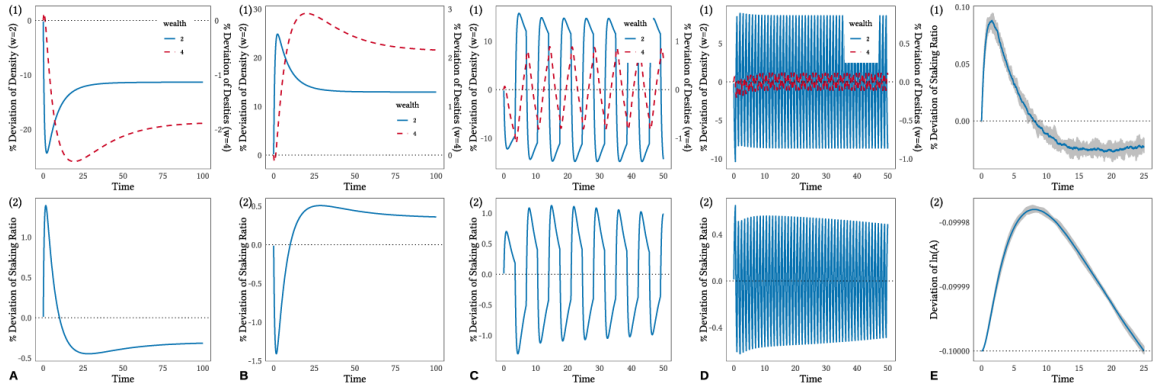


Figure 4: Impulse response paths under different realization of shocks.
Panels A-D show the impulse responses to different shock paths, i.e., continuous good and bad news, low and high-frequency oscillations, respectively. The time unit is roughly calibrated to one-day. In each panel, subplot (1) shows the response of $m_t(w)$, i.e. their deviations from the steady states, with different w. Subplot (2) shows the response of staking ratio. Panel E shows the Monte Carlo simulation of a -0.1 negative shock to $\ln(A)$, where the subplots show the deviations of staking ratio and log productivity. The grey region includes the curves under fewer simulations.

Panel E shows an experiment that involves a negative "MIT shock" to the platform productivity at $t = 0$ and compare it to the steady benchmark. The deviations in the staking ratio are similar to those in Panel B, but the explanation differs. Here, agents increase their staking ratios relative to the benchmark not because they have more wealth,

[32] The system including the wealth distribution will never be stable with continuously arriving shocks (e.g., in panel C, D). In panels A and B, the simulated series only receive shocks in the first week. Nevertheless, we see that these shocks have long-lasting effects to the wealth distribution.

[33] At a specific period, the response can be roughly viewed as some combination of the continuous impacts of historical shocks. Online Appendix OA2 tests the impacts over time of a separate unit shock and also details the method that allows setting up a sequence of shocks.

but because transaction convenience drops. The higher staking ratio provides a larger platform growth drift, which partially offsets the productivity loss caused by the shock. However, when compared to the benchmark, agents continue to experience lower wealth growth, gradually leading to relatively lower staking ratios. Consequently, the productivity falls behind. This indicates that the staking mechanism plays a role in shaping platform growth, yet it is insufficient to fully compensate for significant productivity shocks.

5 Data, Stylized Facts, and Hypotheses

5.1 Data on Stakable Tokens

Our main data source is *Stakingrewards.com*, arguably the largest collector of information related to staking covering both historical and real-time data on most stakable assets. Our sample covers daily observations of 66 stakable tokens including Ethereum 2.0. By the end of 2021, our sample set makes up 37.78% of the total cryptocurrency market capitalization (64.34% when excluding Bitcoin), and respectively consists 80.35% of the PoS market and 97.88% of the DeFi market.[34] The sample period covers July 2018 through Nov 2022, covering the initial birth and rapid growth of "staking," as well as the bear market during 2022. Our sample includes all stakable assets with a market value of more than 100 million US dollars at a snapshot of Aug 2020 (as we initially wrote the paper), some of which have collapsed in 2022. Additional information about staking is typically aggregated from the official websites of each token, including details of staking participation methods (Online Appendix OA1.2), reward sharing rules, real-time staking amount (staking ratio), etc.

Table 1 displays the summary statistics. In most analyses, we also aggregate the daily observations into weekly and monthly data. Panel B shows the large dispersion among tokens in reward rate, staking participation, and price returns: the mean staking reward rate ranges from 0.02% to 75.39%, while the mean staking ratio ranges from 2.78% to 97.77%.

5.2 Stylized Facts about Staking and Token Pricing

Aggregate trends. The shift of focus away from PoW and onto the PoS consensus algorithms have been evident and timely.[35] The PoS share has increased substantially over time from 5% in Oct. 2019 to over 20% in Oct. 2021. As of Oct. 2021, the PoS market cap

[34] According to *CoinMarketCap* and *StakingRewards*. The sample set includes 48 base-layer pan-PoS protocols and 29 high-layer DeFi platform tokens. The two are not mutually exclusive as mentioned.

[35] According to 2021 Staking Ecosystem Report by *StakingRewards*.

Table 1: Summary statistics.
Panel A summarizes the main raw variables. Their token-grouped means and standard deviations are calculated and summarized in Panel B.

Panel A: Raw variables.

	Daily			7-Day			30-Day		
	N	Mean	Std.Dev	N	Mean	Std.Dev	N	Mean	Std.Dev
Reward Rate, r (%, Annual)	41,003	13.42	17.78	5,867	13.28	16.05	1387	13.16	15.69
Reward Ratio, ρ (%)	39,546	6.31	8.86	5,660	6.26	8.58	1339	6.20	8.47
Staking Ratio, Θ (%)	41,706	46.37	23.06	5,964	46.39	23.09	1415	46.36	23.13
Price Appreciation, r_{price} (%)	43,114	-0.04	7.25	6,091	-0.34	22.93	1391	-2.13	54.42
Δr	40,788	-0.03	7.86	5745	-0.13	4.68	1,301	-0.51	8.76
$\Delta\Theta$	41,505	0.00	1.33	5851	0.03	3.37	1,333	0.03	6.25

Panel B: Group-summarized values.

	N	Mean	Std.Dev	Min	25%	Median	75%	Max
$mean(r), (\%)$	66	14.86	15.44	0.02	6.33	9.84	15.77	75.39
$sd(r), (\%)$	66	7.21	10.39	0.01	1.43	2.47	9.60	43.69
$mean(\Theta), (\%)$	65	44.05	22.22	2.78	27.26	44.22	59.61	97.77
$sd(\Theta), (\%)$	65	8.16	5.98	0.27	3.50	6.81	11.69	25.48
$mean(r_{price}), (\%)$	66	-0.01	0.50	-1.17	-0.23	0.04	0.15	2.72
$sd(r_{price}), (\%)$	66	6.76	3.05	3.89	5.53	6.25	6.80	23.88

reached $326.775 Billion with annual growth rate 1,550%, while the overall crypto market cap is up by 673%. The entire staking economy has grown to over 4 million total users by the end of 2021. In recent two years, the staking economy is known to have gone through a full bull/bear cycle along with the cryptocurrency market.

Staking rewards and token price returns. The study of stakable tokens is related to international finance. Token price and staking reward rate can be compared to exchange rates and interest rates. Figure 5 illustrates the excess return in the next week against the interest rate spread calculated as the "foreign interest rate" minus the "local interest rate." Each grey dot represents a weekly data point for a particular token and the blue line shows a linear fit. The upward slope indicates that an increase in the foreign interest rate (relative to the local one) is associated with an increase in the excess return on the token over the local currency, i.e., the crypto version of the "UIP puzzle," which is formally tested and discussed in Section 6.3.

5.3 Hypothesis Formulation

Our model generates a rich set of predictions and reveals mechanisms underlying certain empirical regularities. However, tests on the model implications concerning onchain and offchain wealth distribution, expected steady states, etc., are limited by data availability. We collect data to test the three main sets of model predictions. The first concerns the determination of staking ratio. Proposition 2 implies that **(H1a) a higher staking reward ratio ρ corresponds to a higher staking ratio Θ**. Proposition 5 suggests that

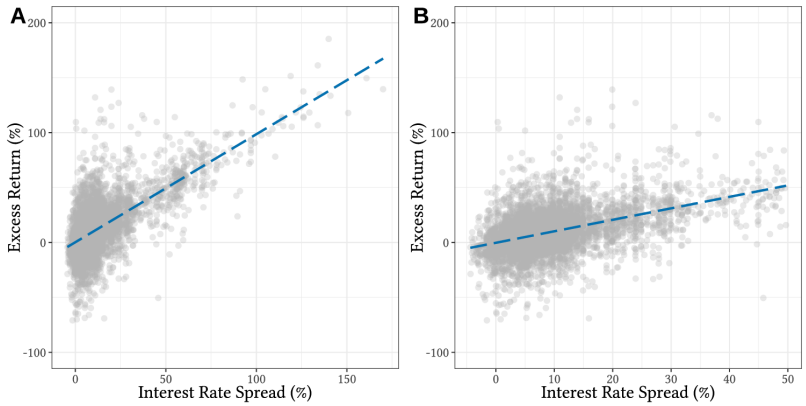


Figure 5: Reward rate spread and token exchange rate.
This figure shows the relationship between the reward (interest) rate spread and the token excess return (exchange rate and interest spread). We treat US dollar as local currency and the 1Y treasury interest (obtained from Fed) rate as the local interest rate. Each dot in the a weekly data point for a particular token. The staking reward rate (annualized) is treated as the foreign interest rate, and the x-axis, the interest rate spread, is calculated as the foreign interest rate minus the local interest rate. The y-axis is the excess return in the next week, i.e., interest rate spread plus the price appreciation. Panel A shows the whole sample set, whereas Panel B limits the interest rate spread to a relatively common range ($< 50\%$). The blue line show linear fitting of the scatter points in each panel.

heterogeneity also impacts equilibrium. In particular, **(H1b) a higher share of large investors is associated with a higher staking ratio**. In practice, it may take time to reach the cross-sectional equilibrium. Therefore, we expect: **(H1c) a higher reward rate r predicts an increase in staking ratio in the short term**. Regarding token prices, Proposition 3 shows a positive predictability of staking ratio on price appreciation, i.e., **(H2a) staking ratio Θ positively predicts token returns**. Finally, concerning crypto carry and interest rate parity, Proposition 4 shows: **(H3a) uncovered interest rate parity is violated among stakable tokens**. Consequently, we expect: **(H3b) crypto carry trade strategies are profitable** and **(H3c) carry predicts excess return in tokens**.

6 Empirical Findings

6.1 Linking Reward Rate and Wealth Concentration to Staking

To test **(H1a)**, we first calculate the daily average of aggregate staking reward ratio $\bar{\rho}$ and staking ratio $\bar{\Theta}$ for each token over its entire sample period. Figure 6 plots their pooling relationship, in which each token generates one dot. The positive-sloped linear fit indicates that the reward is positively related to the staking ratio. This pattern corroborates Proposition 2, and implicitly illustrates that averaging over the time series roughly conforms to the equilibrium. The positive correlation still holds even with a larger slope

after removing high-influential points with $\bar{\rho} > 15\%$. We also test the relationship under earlier data coverage (up to Oct. 2020) in Panel B. Although there are fewer stakable tokens, the significant positive correlation still exists, suggesting the relationship is robust against sample periods.

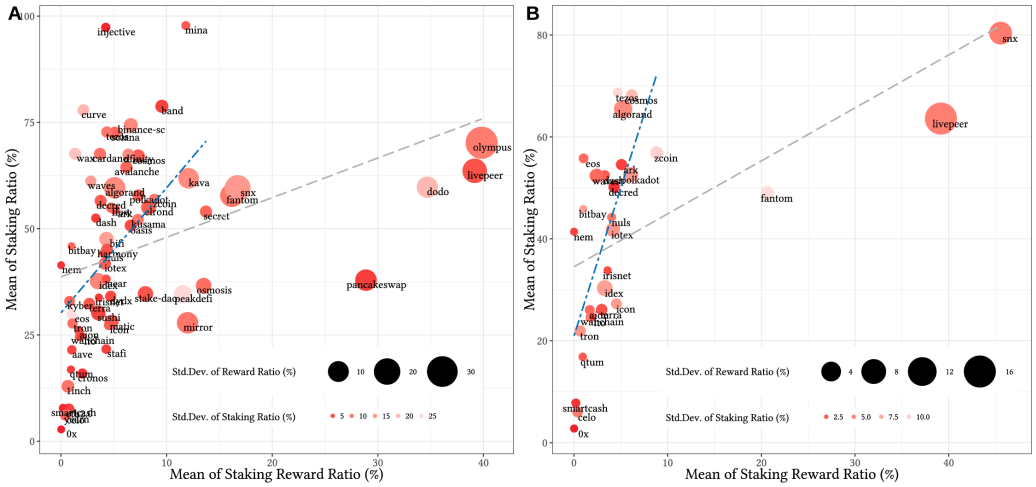


Figure 6: Staking reward versus staking ratio.
This figure corresponds to **(H1a)**, the relationship between staking reward ratio ρ and staking ratio Θ. In Panel A, each token generates one point by calculating its mean Θ and ρ over the sample period. The size and color of the points indicate the standard deviations of ρ and Θ respectively. The gray dashed line is the linear fit. The blue line is the fit after removing the influential points with $\bar{\rho}$ larger than 15%. Panel B plots the same patterns under earlier data coverage (up to Oct. 2020).

We further test **(H1a)** in panel regressions (Table 2). We use the staking ratio of token i at time t, $\Theta_{i,t}$, as the dependent variable, and the staking reward ratio, $\rho_{i,t}$, as the main independent. As Column (1) shows, the estimated coefficient implies a 10-percentage-points higher aggregate reward ratio is associated with a 7.79-percentage-points ($se = 1.90$) higher staking ratio. After inducing time-varying platform controls and investment-related controls, the corresponding estimates decrease to about 3.7 points yet still significant. Monthly sample period and two-way fixed effects are considered. As suggested by Petersen (2008); Abadie et al. (2017), we cluster standard errors on both token and time dimensions together with the fixed effects, which addresses potential heterogeneity in the treatment effects. The positive correlation remains robust. Recall that staking essentially acts as an inflation tax, i.e., facilitates redistribution from more usage-focused adopters to richer or investment-focused adopters. A larger staking reward ratio corresponds to a heavier tax and, forces lower usage preference.

The time-varying token-specific controls bring alternative intuitions in line with our

model: (i) the equilibrium staking ratio negatively relates to the platform productivity, captured by the proxy variable $a_{i,t}$, since it increases transaction convenience with certain staking rewards;[36] (ii) the share of large-asset users (big whales), $whale_{i,t}$, positively relates to staking ratio. This links to heterogeneity and **(H1b)**, and will be discussed specifically later. In addition, token age is a concern as it usually relates to reward designs. Incipient platforms may have not distributed tokens to many agents, thus should be associated with lower staking. We capture age effects by two dummies, $NotLaunched_{i,t}$ and $Y^0_{i,t}$, which equal 1 when the token is in the stage before and within one year of launching on exchanges, respectively. Table 2 corroborates the common hypothesis relates to token age.

Table 2: Staking ratio with respect to the staking reward ratio.

This table tests **(H1a)**, i.e., the relationship between staking ratio, $\Theta_{i,t}$, and the aggregate staking reward ratio, $\rho_{i,t}$, in the same period. Controls include token age, captured by two dummies, $NotLaunched_{i,t}$ and $Y^0_{i,t}$, which equal 1 when the token is before, and within one year of launching on mainstream exchanges, respectively; the proxy of platform productivity, $a_{i,t}$; the value share of tokens held by big-asset accounts, $Whale_{i,t}$, the token market cap, and price volatility. Token and time effects are fixed. Sample sets with different horizons (weekly and monthly) are tested. Standard errors clustered in both token and time dimensions in parentheses. ***,**,* indicate statistical significance at the 1%, 5% and 10% respectively.

Dependent:	$StakingRatio_{i,t}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	7-Day					30-Day			
$\rho_{i,t}$	0.779***	0.428**	0.429**	0.360**	0.374**	0.794***	0.439***	0.413**	0.431**
	(0.190)	(0.162)	(0.161)	(0.170)	(0.173)	(0.188)	(0.155)	(0.155)	(0.173)
$NotLaunched_{i,t}$			-0.076	-0.162**	-0.183**			-0.113**	-0.130**
			(0.064)	(0.063)	(0.064)			(0.053)	(0.051)
$Y^0_{i,t}$			0.004	-0.059	-0.073			-0.028	-0.042
			(0.030)	(0.047)	(0.045)			(0.038)	(0.037)
$a_{i,t}$				-0.747***	-0.417**			-0.623***	-0.415***
				(0.158)	(0.159)			(0.143)	(0.098)
$\frac{1}{100}\log(Cap)_{i,t}$				1.686	1.991			1.124	1.582
				(1.756)	(1.719)			(1.843)	(1.809)
$\frac{1}{100}Volatility_{i,t}$				0.236	0.328			1.590	0.965
				(0.344)	(0.397)			(1.806)	(1.459)
$Whale_{i,t}$					0.217***				0.171**
					(0.071)				(0.075)
Token FE		Yes	Yes	Yes	Yes		Yes	Yes	Yes
Time FE		Yes	Yes	Yes	Yes		Yes	Yes	Yes
Observations	5,660	5,660	5,660	1,364	1,364	1,339	1,339	308	308
R^2	0.088	0.811	0.812	0.917	0.920	0.089	0.801	0.929	0.931

(H1a) focuses on contemporaneous correlations in equilibrium. As mentioned, it takes time to achieve equilibrium in practice, then **(H1c)** expects the reward rate r should pre-

[36]The proxy variable of platform productivity $a_{i,t}$ is measured as the average onchain transaction processing per second. Online Appendix OA3.6 discusses the choice of this proxy.

dict future staking ratio Θ in short terms (transition process). Also, when agents face multiple tokens, the predictability could appear in the cross-section. Table 3 reports the tests for **(H1c)**. We use the change in staking ratio, $\Delta\Theta_{i,t} = \Theta_{i,t} - \Theta_{i,t-1}$, as the dependent variable, and the previous reward rate, $r_{i,t-1}$, as the main independent. The estimated coefficients are all positive, implying that a larger reward rate predicts a positive change in staking ratio, e.g., column (6) shows that if the annual reward rate increases by one percentage point, the staking ratio will increase by 0.026 ($se = 0.015$) percentage points in the following week. This can be a large effect considering the magnitude of the changes in the rewards rate in the staking economy. Previous staking ratio $\Theta_{i,t-1}$ is controlled to capture the potential diminishing marginal effects. Time-varying platform controls are also considered, which do not exhibit significant influences. This is reasonable since the platform characteristics enter the equilibrium, but their impact on the transition process could be complex. In addition, the statistical predict power decreases as the time window expands. This is partly due to the accumulation of noise over a longer time period. More importantly, the longer period makes the predicted impact already reflected in the formation of new equilibrium.

Wealth concentration and aggregate staking. As Section 4 illustrates, agents are heterogeneous in staking propensities. In particular, large-asset agents (whales) tend to stake more, implying that when the tokens are more concentrated in whales, the equilibrium staking ratio would be higher **(H1b)**. We test this hypothesis in columns (5) and (9) of Table 2, where the coefficients of $whale_{i,t}$ are significantly positive. That is, concentration implies additional general tendencies to staking. Importantly, if staking benefits the platform in a fundamental way (such as enhanced network security), wealth concentration may not be all bad — large stakeholders are more interested in becoming facilitators of the platform's services than pure users. Agents, on the other hand, may have incentives to pursue larger usage benefits at the cost of more concentration, as discussed earlier. In practice, staking/voting pools for prominent platforms are often backed by large exchanges and foundations, which is consistent with the findings here.

6.2 Equilibrium Staking Ratio and Token Price Dynamics

Table 4 reports the tests for **(H2a)**, i.e., staking ratio positively predicts token price changes. We calculate the log token price change, $r_{price_{i,t}} = \log(\frac{P_{i,t}}{P_{i,t-1}})$, and regress it on the previous staking ratio. We consider the cryptocurrency-market and token-size factors that

Table 3: Staking ratio with respect to the staking reward rate.

This table tests **(H1c)**, i.e., how people's staking choices are affected by the reward rate. The dependent is the change of staking ratio, $\Delta StakingRatio_{i,t}$, and the main independent is the reward rate in the previous period, $r_{i,t-1}$. Controls include token age, proxy of previous platform productivity, previous percentage of tokens held by big-asset accounts that are similar in previous tables, and previous staking ratio, $StakingRatio_{i,t-1}$ that capturing potential diminishing marginal effect. Token and time effects are fixed. Sample set with different horizons (weekly and monthly) are tested. Standard errors clustered in both token and time dimensions in parentheses. ***,**,* indicate statistical significance at the 1%, 5% and 10% respectively.

Dependent:	$\Delta StakingRatio_{i,t}$								
	Daily					7-Day		30-Day	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$r_{i,t-1}$	0.002***	0.003***	0.001**	0.001**	0.002*	0.026*	0.030	0.049**	0.006
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.015)	(0.024)	(0.020)	(0.019)
$StakingRatio_{i,t-1}$			-0.009***	-0.009***	-0.016***		-0.103***		-0.294***
			(0.002)	(0.002)	(0.004)		(0.028)		(0.057)
$NotLaunched_{i,t}$				-0.001	-0.002*		-0.009		-0.029
				(0.001)	(0.001)		(0.006)		(0.021)
$Y^0_{i,t}$				0.000	-0.001		-0.006		-0.012
				(0.000)	(0.001)		(0.005)		(0.016)
$a_{i,t-1}$					0.005		0.018		0.046
					(0.006)		(0.047)		(0.107)
$Whale_{i,t-1}$					0.003		0.019		-0.002
					(0.002)		(0.011)		(0.036)
Token FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39,359	39,359	39,359	39,359	10,636	5,559	1,511	1,266	344
R^2	0.0006	0.043	0.049	0.049	0.172	0.063	0.197	0.124	0.322

have important impacts on price changes as discussed in Liu et al. (2019). The estimated coefficient of staking ratio is significantly positive, indicating that a higher staking ratio predicts larger token price appreciation, e.g., column (4) shows that if the staking ratio of a token increases by one percentage point, its price will appreciate by 0.066% ($se = 0.023\%$) in the next week. Recall the large variation in the staking ratio, this effect can have a large impact on price. This effect is robust against different time windows. Staking ratio has incremental predictive power with more control variables that have been suggested to affect the token's pricing in recent studies, including platform controls (token age, productivity, and whale share) for possible predictable price appreciation from platform characteristics, the previous returns for the potential momentum, and onchain network effects, $\Delta Network_{i,t-1}$.[37] To further consider the time-series auto-correlation, we adopt two-way fixed effects regression in columns (3), (6), and (9), where the estimated rolling

[37]It is measured by the lagged log differences in the total amount of addresses with non-zero balance on the platform. As Cong et al. (2021b) discusses, cryptocurrency returns exhibit network adoption premia. The estimated coefficient of the network adoption term is positive and consistent with prior research.

CAPM β captures the forces from market fluctuations. The estimated coefficients remain significant and positive with two-way clustered standard errors.

Table 4: Staking ratio and token price.

This table tests **(H2a)**, i.e. staking ratio predicts token price appreciation. The dependent $r_{price_{i,t}}$ is the log price change. The main independent is the staking ratio of the previous period, $StakingRatio_{i,t-1}$. Controls include the market return MKT_t, the log market cap $\log(Cap)_{i,t-1}$, the proxy of network adoption $\Delta Network_{i,t-1}$, the price return of the previous period $r_{price_{i,t-1}}$, and rolling CAPM beta, $\hat{\beta}_{i,t}$. Token characteristic controls include the token age, platform productivity, capital share held by big-asset accounts. We also do the test in different time spans and with token-specific and time fixed effects. Standard errors clustered in both token-specific and time dimensions in parentheses. ***,**,* indicate statistical significance at the 1%, 5% and 10% respectively.

Dependent:	$r_{price_{i,t}}$								
	Daily			7-Day			30-Day		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$StakingRatio_{i,t-1}$	0.009**	0.027***	0.022**	0.066***	0.172**	0.138*	0.208*	0.347*	0.372**
	(0.004)	(0.007)	(0.008)	(0.023)	(0.068)	(0.071)	(0.121)	(0.197)	(0.139)
MKT_t	0.968***	1.029***		0.844***	0.685*		2.445*	2.201	
	(0.031)	(0.043)		(0.264)	(0.352)		(1.435)	(1.496)	
$\hat{\beta}_{i,t}$			-0.002			-0.037			-0.132
			(0.002)			(0.031)			(0.104)
$\log(Cap)_{i,t-1}$	-0.002***	-0.002**	-0.005***	-0.027***	-0.031***	-0.038***	-0.120***	-0.121***	-0.113***
	(0.000)	(0.001)	(0.001)	(0.006)	(0.009)	(0.009)	(0.034)	(0.043)	(0.021)
$r_{price_{i,t-1}}$		0.021	0.035		0.008	-0.075*		0.127*	-0.076
		(0.050)	(0.060)		(0.040)	(0.042)		(0.062)	(0.074)
$\Delta Network_{i,t-1}$		0.167***	0.224***		0.195	0.366		0.992	0.996
		(0.058)	(0.068)		(0.207)	(0.259)		(1.393)	(1.216)
$a_{i,t-1}$		0.047	0.069		0.603**	0.306		1.007	0.614
		(0.030)	(0.041)		(0.258)	(0.235)		(0.825)	(0.946)
$Whale_{i,t-1}$		-0.010	-0.013		-0.006	-0.103		-0.179	-0.253
		(0.009)	(0.009)		(0.086)	(0.073)		(0.341)	(0.201)
$NotLaunched_{i,t}$		-0.003	0.011***		0.075***	0.108**		0.119	0.159
		(0.002)	(0.004)		(0.024)	(0.040)		(0.154)	(0.126)
$Y^0_{i,t}$		0.002	0.007**		0.021	0.056**		-0.073	0.114
		(0.002)	(0.003)		(0.021)	(0.020)		(0.084)	(0.107)
Token FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE			Yes			Yes			Yes
Observations	41,544	10,887	9,991	5,872	1,530	1,434	1,347	334	322
R^2	0.267	0.346	0.478	0.043	0.054	0.507	0.120	0.207	0.640

Further discussion may lie in the heterogeneous predictive power of staking ratio across tokens categories and market sentiment. We test the same specification on multiple sub-samples, including bull/bear sub-periods, pan-PoS/DeFi sub-samples in Online Appendix OA3.6 and OA3.7. The estimated coefficients of $\Theta_{i,t-1}$ are all positive, among which only the bear-period exhibits lower statistical significance. This suggests the positive relationship between staking ratio and expected price appreciation to be a generally existing phenomenon in our sample. As applications to such predictability, we also ex-

amine the performance of Θ-sorting portfolio strategies in Online Appendix OA3.8.

6.3 UIP Violation and Crypto Carry

UIP violation. We test **(H3a)**, i.e., whether UIP is violated, using the specification in Fama (1984): define the excess return of token i at t as $\lambda_{i,t} = \log P_{i,t+1} - \log P_{i,t} + (r_{i,t} - c_i) - r_t^f$, and regress $\lambda_{i,t+1}$ on $r_t^f - r_{i,t} + c_i$ with coefficient B and token fixed effects,[38] where P_t is the price in local currency, r^f is the local interest. Under UIP, $B = 0$, i.e. the excess return is not forecastable by current interest rate differences. We examine different time horizons as Valchev (2020) does, and use different assets as local, including US dollar, Bitcoin, Ethereum and stakable ETH 2.0. Table 5 reports the findings. All the results show significantly negative estimated B, implying that a higher interest rate will predict a positive appreciation of the exchange rate. This leads to potential arbitrage opportunities. We also use each single token as local currency, respectively. The regression results (reported in Online Appendix OA3.9) suggest the violation of UIP also exists within the token market.

Table 5: Test of UIP violation.

This table tests **(H3a)**, i.e., regressing excess returns on previous reward rate spread (with coefficient B) with token-specific effects. In each row, we use a different asset as local currency and report the estimated coefficients and standard errors (clustered by tokens) of B.

Local	7-day			30-day		
Currency	Coef., B	Std. Err.	R^2	Coef., B	Std. Err.	R^2
US Dollar	−1.02	(0.044)	0.33	−1.12	(0.176)	0.11
Bitcoin	−1.02	(0.034)	0.37	−1.08	(0.134)	0.11
Ethereum	−1.04	(0.033)	0.37	−1.09	(0.126)	0.12
Eth 2.0	−1.04	(0.014)	0.42	−1.25	(0.059)	0.17

Crypto carry trades. UIP violations naturally prompt us to examine the predictability of crypto carry to token excess return and the performance of the crypto carry trade portfolio **(H3b)**. Tokens in the asset pool are ordered by their carry in the previous period, and then divided into the top 50% and the bottom 50%. A carry trade portfolio is constructed by going long high-carry group with equal weight and going short low with equal weight at the end of each week. Long tokens are also staked to earn staking reward rates, while short tokens pays additional compensate for the staking reward rate. The portfolio is rebalanced every week.[39]

[38] While in practice, there are various ways to stake (e.g., delegating and running a node), which corresponds to different reward rates and costs, the staking programs mostly feature delegation/voting (that our data correspond to), which incurs negligible time-varying operational costs. We therefore assume $c_{\{i,t\}}$ for each token is constant (c_i) and eliminated by token-specific fixed effects.

[39] Most stakable tokens offer the flexibility of weekly staking or have derivatives that enable such an asset

Table 6 reports the statistics of these strategies. The carry strategy has a significantly greater positive return and yields a Sharpe ratio of 1.60. Since we long high carry and short low carry, the portfolio carry is always positive. If the portfolio always achieves positive returns, it means that in the cross-section, assets with higher carry have greater aggregate returns. For higher moments, the strong positive skewness is associated with the currency carry trade shown by Brunnermeier et al. (2008). Moreover, the carry strategy exhibits excess kurtosis, indicating fat-tailed positive and negative returns, which is consistent with Koijen et al. (2018)'s findings for currencies and commodities. The long-short carry trade strategy exhibits relatively stable returns, especially considering the high volatility of cryptocurrency markets and the bull/bear circle during 2020-2022.

We also report related strategies for comparison: (i) hold the same portfolio but without stake (and not compensate for staking); (ii) apply the same strategy but rebalance every month; (iii) considering the potential short-selling limits, examine the average return of equal-weighted full-sample / (top 50%) high-carry / low-carry tokens. The non-staking strategy also yields positive returns, implying that the excess returns are not only from carry (staking reward) also price appreciation. Moreover, the monthly-rebalanced strategy exhibits fewer returns. There are two explanations: (i), the reward rate decreases with contemporaneous staking ratio mechanically. Therefore, investors cannot consistently earn high carry over a long period without timely position adjustments; (ii), the reversal of reward rate further influences the staking ratio, which then weakens the effect on price appreciation. Finally, long-only strategies corroborate the carry premia: longing top 50% tokens with high carry outperforms the simple equal-weighted benchmark, while the bottom 50% performs the worst. Their cumulative returns plotted in Online Appendix OA3.9 ensures the above observations.

Excess return predicted by carry. In **(H3c)**, the return predictability can come from both the crypto carry itself and any price appreciation that is related to or predicted by carry. We follow Koijen et al. (2018) to regress the overall excess return on the previous carry. Table 7 reports the estimations of the coefficient of carry, C. The results indicate that carry is a strong predictor of expected return. Without token-specific fixed effect, the estimated coefficient is around 1, which means that high staking reward rate tokens neither relatively depreciate nor appreciate on average. Hence, investors can earn reward

allocation. Considering the abnormal fluctuation when a staking project first launches, our weekly asset pool does not include new staking projects that appear within the past week.

Table 6: Statistics of carry strategies.

This table corresponds to **(H3b)** and reports the statistics of carry trade strategies. The first three rows report the results of the long-short carry strategies. The rows below report long strategies. Annualized mean, standard deviations, skewness, kurtosis, maximum drawdown (MDD), and Sharpe ratio are reported.

Strategy	Mean (Annual, %)	St.dev. (Annual, %)	Skewness	Kurtosis	MDD (%)	Sharpe Ratio (Annual)
Long-short Strategy:						
1W-Carry Trade (Staking)	65.820	41.095	1.410	18.772	29.966	1.602
1W-Carry Trade (Non-staking)	52.497	41.104	1.404	18.719	35.920	1.277
1M-Carry Trade (Staking)	45.051	56.929	1.260	20.508	69.497	0.791
Long Strategy:						
EW All assets	15.577	78.244	−1.576	7.672	92.934	0.199
EW High Carry	49.416	81.309	−1.103	4.687	90.419	0.608
EW Low Carry	−16.404	80.444	−1.804	9.907	95.645	−0.204

rate spread using carry trade. With token fixed effect, the estimated C is less than 1 and even insignificant in monthly specifications.[40] This suggests time series carry predicts less expected return due to the staking mechanism:[41] despite higher reward rates lead to higher staking ratios and further price appreciations, there is a *downward adjustment effect* of the reward rate in the time series automatically adjusted for the increment of staking ratios. Then excess return is lowered by the adjustment. The downward adjustment is magnified when the time window expands, and thus the estimated C in columns (6), (8) are smaller than (2) and (4). This reflects the consequence of the staking reward rate being determined by the competitive equilibrium.

Table 7: Carry and excess returns.

This table tests **(H3c)**. The dependent variable is the excess return, and the independent is the carry in the previous period. Standard errors clustered in both token-specific and time dimensions in parentheses. ***,**,* indicate statistical significance at the 1%, 5% and 10% respectively.

Dependent:	$\text{ExcessReturn}_{i,t}$							
	7-Day				30-Day			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{Carry}_{i,t-1}$	0.956***	0.901***	0.968***	0.917***	0.968***	0.773	1.009***	0.846**
	(0.053)	(0.095)	(0.042)	(0.071)	(0.296)	(0.534)	(0.216)	(0.383)
Token FE		Yes		Yes		Yes		Yes
Time FE			Yes	Yes			Yes	Yes
Observations	5,745	5,745	5,745	5,745	1,301	1,301	1,301	1,301
R^2	0.230	0.239	0.441	0.447	0.038	0.094	0.333	0.374

[40] Without token-specific and time-fixed effects, C represents the total predictability of returns from carry from both its passive and dynamic components. Token fixed effects will remove the predictable return component of carry that comes from passive exposure to tokens with different unconditional average returns.

[41] This phenomenon is similarly found in commodities (Koijen et al., 2018) with a different mechanism: when a commodity has a high spot price relative to its futures price, implying a high carry, the spot price tends to depreciate on average, thus lowering the realized return on average below the carry.

7 Conclusion

Staking has become a hallmark feature in many distributed networks involving hundreds of billions of dollars. In addition to offering a convenience yield for transactions, blockchain-based tokens are frequently staked for base-layer consensus generation or for incentivizing economic activities in DeFi protocols and platform development, and consequently earn staker's rewards. We build the first dynamic model of a token-based economy where agents endogenously allocate wealth on and off a digital platform and use tokens either to earn rewards or to transact. We solve this mean-field game with stochastic controls and systematic shocks, and identify staking ratio as a fundamental variable linking staking to the endogenous reward rate and token price. The staking ratio is proportional to the reward rates in the cross-section but negatively correlated to reward rates in the time series; it positively predicts the returns of cryptocurrencies. Furthermore, the model rationalizes violations of the uncovered interest rate parity, and the significant crypto carry premia that we empirically document.

The framework can be explored further for studying the utilities of platform tokens. For example, DeFi projects increasingly lock up both native and non-native tokens. Allowing multiple tokens to be used within a network may cause the payment utility of native tokens to decline. But stakable tokens entitle the holders to instead collect rewards (fees and subsidies), while providing functionalities such as security or liquidity for the networks. Given that many platforms use staking to foster adoption and demand, optimally designing the various utilities of tokens and understanding their implications on token prices constitute interesting future research. Similarly, it remains an open question how to jointly design token supply policy and staking protocols.

Appendix

Proof of Proposition 1. The marginal utility of staked and non-staked tokens are

$$MU_l = MU_o + (r_t - c_t)\frac{\partial J}{\partial w}, MU_x = MU_o + (1-\alpha)\left(\frac{N_t A_t U_t}{x_t}\right)^{\alpha}\frac{\partial J}{\partial w}, \tag{A.1}$$

where $MU_o = (\mu_t + \frac{\partial \Psi}{\partial n_t})\frac{\partial J}{\partial w} + q_t\sigma_t^2\frac{\partial^2 J}{\partial w^2}$, $q_t = x_t + l_t$. is the common part of MU_l and MU_x. Note that $MU_l - MU_o$ stays the same in l, while $MU_x - MU_o$ decreases in x. Thus, when treat q_t, is given, there exists at most a unique $\tilde{x}_t \leq q_t$ that satisfies $MU_l = MU_x$, i.e.,

$\tilde{x}_t = \left(\frac{1-\alpha}{r_t - c_t}\right)^{\frac{1}{\alpha}} N_t A_t U_t$.[42] Substitute (9) and note the boundaries, then (15) is obtained.

Proof of Proposition 2. When there is an inner point solution of individual staking ratio, the on-chain allocation is already in optimal, i.e., $MU_l = MU_x$. Then the marginal onchain utility in optimal reads $MU_q = (\mu + r - c + \partial_n \Psi)\partial_w J + q\sigma^2 \partial_{ww} J$, where the subscript of time t is omitted. By risk aversion and the definition of Ψ_t, MU_q decreases in q. Note that under a risk-aversion setting, when $q \to \infty$, $MU_q \to -\infty$. Thus, there exists a unique $q^* \geq 0$ that satisfies the first-order condition, $MU_q = 0$, if $MU_q(q = 0) > 0$. Otherwise, $q^* = 0$.[43] Consider the case when $q = x$. Similarly, the marginal transaction convenience is decreasing, leading to decreasing MU_q and a unique solution to the F.O.C.

In equilibrium, $\int_i x_i^* \mathrm{d}i = \left(\frac{1-\alpha}{r_t - c_t}\right)^{\frac{1}{\alpha}} N_t A_t \int_i U_{it} \mathrm{d}i = \left(\frac{1-\alpha}{r_t - c_t}\right)^{\frac{1}{\alpha}} N_t A_t U_t$.[44] Then $\Theta^* = 1 - \left(\frac{1-\alpha}{r_t - c_t}\right)^{\frac{1}{\alpha}} \frac{N_t A_t U_t}{q_t^*} \equiv \Theta^*(r_t)$. r_t^* fits (11). Substituting into agents' optimization, we obtain

$$0 = (1-\alpha)\left[\frac{N_t(\Theta_t^*)A_t U_t}{q_t^*(1-\Theta_t^*)}\right]^{\alpha} - \frac{\rho}{\Theta^*} + c_t \equiv h_1(\Theta^*(\rho))q^*(\rho)^{-\alpha} - h_2(\rho). \tag{A.2}$$

Clearly, both Θ^* and q^* are affected by the reward ratio ρ given A and w.

$$\Theta'(\rho) = \frac{\Theta^* + \Theta^{*2}\alpha h_1(\Theta^*(\rho))q^*(\rho)^{-\alpha-1}q^{*\prime}(\rho)}{\Theta^{*2}h_1'(\Theta^*(\rho))q^*(\rho)^{-\alpha} + \rho}, \tag{A.3}$$

where the denominator is positive. Consider $q^{*\prime}$. $\frac{\partial MU_q}{\partial \rho} = (1-\Theta)\partial_w J + \sigma^2 \partial_{ww} J q^{*\prime}(\rho)$. Suppose there is an ρ, s.t. $q^{*\prime}(\rho) < 0$. Under risk-aversion, $\partial_{ww} J < 0$ and thus $\frac{\partial MU_q}{\partial \rho} > 0$ for all q. Since q^* is the solution of F.O.C., $q^{*\prime}(\rho) > 0$, leading to contradiction. Back to (A.3), we obtain $\Theta'(\rho) > 0$. Thus if there is a $\Theta(\rho)$ satisfies (A.2), it must be unique. For existence: the first term on the R.H.S. of (A.2) tends to infinite when $\Theta \to 1$, thus the R.H.S. must be positive. $\Theta \to 0$, N_t tends to zero and thus the R.H.S. must be negative.

Proof of Proposition 3. Substituting the derivatives w.r.t. $V(A_t)$ into (20), we obtain

$$0 = V'(A_t)A_t\mu^A(\Theta_t) + [r(\Theta_t) - c_t - E(\Theta_t, A_t) + \Psi_n(A_t)]V(A_t) - [\sigma^S V(A_t)]^2. \tag{A.4}$$

Then, we show that under current specific forms assumed in Proposition 3, I and Θ_t are both at most functions of A_t and $V(A_t)$, but do not contain any derivatives of A_t. Precisely, $\Psi_t = \psi(A_t)(w_t - n_t)$. From the first-order derivatives of the HJB equation, we obtain

[42] The only case where such $\tilde{x}_t$ does not exist is that $MU_l \leq MU_x$ even when $l_t = 0$. Then, the individual staking ratio is zero. The entry cost in (3) avoids the case where l and x both tend to zero.

[43] Note that q^* could be larger than individual wealth since self-financing is allowed.

[44] We focus on the interior case, since the corner solution leads to a trivial equilibrium of $\Theta^* = 0$ and $\Theta^{*\prime}(\rho) = 0$. The integral seems unnecessary under a representative-agent setting; it, however, helps understand the aggregation of on-chain tradable allocation in the general case with heterogeneity. Also note that $\int_i U_{it} \mathrm{d}i$ equals to U_t since agents are homogeneous and in a unit measure.

$\partial_w J = \mathcal{U}'(y_t) = \frac{1}{y_t}$. Agents with log-utilities face changing investment opportunities, $I = -\partial_{ww} J / \partial_w J = \frac{1}{w_t}$, and the on/off-chain allocation is linear in agent's wealth.[45] Meanwhile, (16) implies that Θ_t satisfies $\rho(\Theta_t, A_t) = \Theta_t \left\{ (1-\alpha) \left[\frac{N(\Theta_t) A_t U_t}{V(A_t)(1-\Theta_t)} \right]^{\alpha} + c_t \right\}$ and thus uniquely exists as a function of A_t and $V(A_t)$ without any derivatives. Therefore, we obtain (21). It follows the form $V'(A) = F(A, V(A))$, where F is continuously differentiable. Therefore, the ODE has a unique solution given the boundary condition.

Proof of Proposition 4. Rearrange the formula of MU_q to obtain (4).

Proof of Proposition 5. Similar to previous proofs, there exists a unique $\tilde{x}_{i,t}$ s.t. $MU_l = MU_x$.[46] Based on the proof of Proposition 2, we denote the unique solution of $MU_q = 0$ as $q^*_{i,t}$. Agents may vary in trade-offs between convenience benefits and the fixed entry cost φ. When $\tilde{x}_{i,t} < q^*_{i,t}$, (3) yields $\mathrm{d}v(\tilde{x}) = \left[\left(\frac{1-\alpha}{r_t - c_t} \right)^{\frac{1-\alpha}{\alpha}} N_t A_t U(u_{i,t}, w_{i,t}) - \varphi \right] \mathrm{d}t$. Only when $U(u_{i,t}, w_{i,t})$ are large enough s.t. $\mathrm{d}v_t > 0$, i.e., $U_t > \frac{\varphi}{NA} \left(\frac{r-c}{1-\alpha} \right)^{\frac{1-\alpha}{\alpha}} \equiv U_0$, transaction is applicable. Otherwise, the optimal θ_{it} is always 1. Then (26) is obtained.[47]

Proof of Proposition 6.

Lemma 1. ***The penalty function.*** *The penalty function $v(w)$ satisfies*

$$v(w) = -\int \frac{\delta J(w, w', m)}{\delta m} \frac{\partial g(w') m(w')}{\partial w'} \mathrm{d}w'. \tag{A.5}$$

Proof. Treat J_t as $J(\cdot, m_t)$, and use Itô's Lemma, $\mathrm{d}J_t = \langle \frac{\delta J}{\delta m}, \mathrm{d}m_t \rangle + \frac{1}{2} \langle \mathrm{d}m_t | \frac{\delta^2 J}{\delta m^2} | \mathrm{d}m_t \rangle$. □

(A.5) always holds even the non-stationary states. Consider the value function.

Lemma 2. *With log utility, the value function J is separable: $J = \frac{\log(w)}{\phi} + F(m, A)$, where F satisfies $\forall w, w', \forall w''$, $\frac{\delta F}{\delta m}(w, w'', m, A) = \frac{\delta F}{\delta m}(w', w'', m, A) \equiv \hat{F}(w'', m, A)$.*

Proof. Consider $J(w, m, A) = b \log(w) + F(m, A)$ as a trial form. Then the optimal $y = 1/\partial_w J = \frac{w}{b}$. $q = \frac{\mu + r - c + \Psi_n}{\sigma^2} w \equiv \frac{k_1}{\sigma^2} w$, where k_1 could be a functional of m and A, but definitely does not involve w, since the endogenous variables, μ and r, are determined by the whole

[45] Proof of Proposition 6 precisely proves this in a more general (heterogeneous) setting. The general solution can be found in Merton (1971) and Duffie (2010). Because the main results do not depend on risk-aversion, we assume log utility for simplicity. For general utilities such as HARA, agents have extra terms to hedge the changes of investment opportunity. The numerical approach still applies.

[46] Note that $\tilde{x}$ reflects the trade-off threshold of switching to staking, while both the two on-chain allocations need to satisfy $MU_q \geq 0$. $\tilde{x}_{i,t} = \left(\frac{1-\alpha}{r_t - c_t} \right)^{\frac{1}{\alpha}} N_t A_t U(u_{i,t}, w_{i,t})$.

[47] Similarly, Online Appendix OA3.3 proves other heterogeneity implications. The key observation is that heterogeneity changes the proofs by entering m, while the overall variables are integrated values over m, then the monotonicity of the staking ratio precludes a great deal of concern about uniqueness.

crowd rather than any atomic agent. $x = \left(\frac{1-\alpha}{r-c}\right)^{\frac{1}{\alpha}} NAu_i w$. Substituting these optimal controls and the derivatives of J into the HJB equation, we obtain

$$\begin{aligned}\phi b \log(w) + \phi F = \log(w) - \log(b) + f(w)\frac{b}{w} - \frac{1}{2}g(w)^2\frac{b}{w^2} + \mu^A A\partial_A F \\ + \frac{\partial}{\partial w}\left(g(w)\int \frac{\delta F}{\delta m}(w, w', m)\frac{\partial g(w')m(w')}{\partial w'}\mathrm{d}w'\right),\end{aligned} \tag{A.6}$$

where $g(w) = \frac{k_1^2}{\sigma}w$, $f(w) = \frac{k_1^2}{\sigma^2}w - \frac{w}{b} + \alpha\left(\frac{1-\alpha}{r-c}\right)^{\frac{1-\alpha}{\alpha}} NAu_i w \equiv (\frac{k_1^2}{\sigma^2} - \frac{1}{b} + k_{2i})w$. By the property of F, the last term can be rearranged as $\frac{k_1}{\sigma}\tilde{F}(m, A)$. Since (A.6) holds for all w, the coefficient of $\log(w)$ should be zero, $b = \frac{1}{\phi}$ and satisfies the supposed form of b. Then

$$\phi F = \log(\phi) + \frac{k_1^2}{2\phi\sigma^2} - 1 + \frac{k_{2i}}{\phi} + \mu^A A\partial_A F + \frac{k_1\tilde{F}}{\sigma}, \tag{A.7}$$

indicating that F does not involve w,[48] i.e., the separable trial form is indeed valid. □

With the lemmas in hand, we consider the steady-state law of motion,

$$0 = L^*(w)[m] = -\frac{\partial}{\partial w}\left(\widehat{f(w,m)}m_t(w)\right) + \frac{\partial^2}{\partial w^2}\left(\frac{1}{2}g^2(w,m)m_t(w)\right), \tag{A.8}$$

where $\widehat{f(w,m)}$ is the average f_i among different types $u_{i,t}$ whose wealth w_i equals to w:

$$\widehat{f(w,m)} = \mathbb{E}[f_i(w_i, m)|w_i = w] = \left(\frac{k_1^2}{\sigma^2} - \frac{1}{b} + \mathbb{E}(k_{2i})\right)w = (\frac{k_1^2}{\sigma^2} - \frac{1}{b} + k_2 w^{\beta-1})w, \tag{A.9}$$

where $k_2 = \alpha\left(\frac{1-\alpha}{r-c}\right)^{\frac{1-\alpha}{\alpha}} NA\bar{u}$. Rearrange the law of motion (A.8), we obtain

$$\left(\frac{k_1^2}{\sigma^2} - \phi\right)\frac{\partial}{\partial w}[wm(w)] + k_2\frac{\partial}{\partial w}[w^\beta m(w)] = \frac{k_1^2}{2\sigma^2}\frac{\partial^2}{\partial w^2}[w^2 m(w)]. \tag{A.10}$$

Let $h(w) = w^2 m(w)$. We obtain a 2nd-order ODE w.r.t. h (let $a_1 = 2 - \frac{2\phi\sigma^2}{k_1^2}$, $a_2 = \frac{2k_2\sigma^2}{k_1^2}$):

$$h''(w) = a_1\frac{h'(w)w - h(w)}{w^2} + a_2\frac{h'(w)w - (2-\beta)h(w)}{w^{3-\beta}}. \tag{A.11}$$

Solve this ODE and note m is a p.d.f. with bounded integral, we obtain h and further (27).

References

Abadi, Joseph and Markus Brunnermeier (2018): "Blockchain economics," Tech. rep., National Bureau of Economic Research.

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge (2017): "When should you adjust standard errors for clustering?" Tech. rep., National Bureau of Economic Research.

[48] The precise solution to $F(m, A)$ needs the determination of how k_1, k_{2i}, μ^A, and σ depends on m and A. This further requires the combination with the pricing formula. However, whatever the precise solution is, (A.7) already indicates that F does not involve w.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2022): "Income and wealth distribution in macroeconomics: A continuous-time approach," The review of economic studies, 89 (1), 45–86.

Augustin, Patrick, Roy Chen-Zhang, and Donghwa Shin (2022): "Yield Farming," Available at SSRN 4063228.

Bailey, Warren and Kalok C Chan (1993): "Macroeconomic influences and the variability of the commodity futures basis," The Journal of Finance, 48 (2), 555–573.

Bakshi, Gurdip and George Panayotov (2013): "Predictability of currency carry trades and asset pricing implications," Journal of financial economics, 110 (1), 139–163.

Bansal, Ravi (1997): "An exploration of the forward premium puzzle in currency markets," The Review of Financial Studies, 10 (2), 369–403.

Bansal, Ravi and Wilbur John Coleman (1996): "A monetary explanation of the equity premium, term premium, and risk-free rate puzzles," Journal of Political Economy, 104 (6), 1135–1171.

Barr, David G and Richard Priestley (2004): "Expected returns, risk and the integration of international bond markets," Journal of International money and finance, 23 (1), 71–97.

Basu, Soumya, David Easley, Maureen O'Hara, and Emin Sirer (2019): "StableFees: A Predictable Fee Market for Cryptocurrencie," Available at SSRN 3318327.

Bech, Morten L and Rodney Garratt (2017): "Central bank cryptocurrencies," BIS Quarterly Review September.

Bekaert, Geert (1996): "The time variation of risk and return in foreign exchange markets: A general equilibrium perspective," The Review of Financial Studies, 9 (2), 427–470.

Benhaim, Alon, Brett Hemenway Falk, and Gerry Tsoukalas (2021): "Scaling Blockchains: Can Elected Committees Help?" arXiv preprint arXiv:2110.08673.

Bertucci, Charles, Louis Bertucci, Jean-Michel Lasry, and Pierre-Louis Lions (2020): "Mean field game approach to bitcoin mining," arXiv preprint arXiv:2004.08167.

Biais, Bruno, Christophe Bisiere, Matthieu Bouvard, and Catherine Casamatta (2019): "The blockchain folk theorem," Review of Financial Studies, 32 (5), 1662–1715.

Biais, Bruno, Christophe Bisiere, Matthieu Bouvard, Catherine Casamatta, and Albert J Menkveld (2020): "Equilibrium bitcoin pricing," Available at SSRN 3261063.

Bilal, Adrien (2023): "Solving heterogeneous agent models with the master equation," Tech. rep., National Bureau of Economic Research.

Brunnermeier, Markus K, Stefan Nagel, and Lasse H Pedersen (2008): "Carry trades and currency crashes," NBER macroeconomics annual, 23 (1), 313–348.

Burnside, Craig, Martin Eichenbaum, Isaac Kleshchelski, and Sergio Rebelo (2011): "Do peso problems explain the returns to the carry trade?" The Review of Financial Studies, 24 (3), 853–891.

Capponi, Agostino and Ruizhe Jia (2021): "The Adoption of Blockchain-based Decentralized Exchanges: A Market Microstructure Analysis of the Automated Market Maker," Available at SSRN 3805095.

Cardaliaguet, Pierre, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions (2019): "The master equation and the convergence problem in mean field games," in The Master Equation and the Convergence Problem in Mean Field Games, Princeton University Press.

Casassus, Jaime and Pierre Collin-Dufresne (2005): "Stochastic convenience yield implied

from commodity futures and interest rates," The Journal of Finance, 60 (5), 2283–2331.

Chiu, Jonathan, Seyed Mohammadreza Davoodalhosseini, Janet Hua Jiang, and Yu Zhu (2019): "Bank market power and central bank digital currency: Theory and quantitative assessment," Available at SSRN 3331135.

Christin, Nicholas, Bryan Routledge, Kyle Soska, and Ariel Zetlin-Jones (2023): "The crypto carry trade," Working Paper.

Cong, Lin William and Zhiguo He (2019): "Blockchain disruption and smart contracts," Review of Financial Studies, 32 (5), 1754–1797.

Cong, Lin William, Zhiguo He, and Jiasun Li (2021a): "Decentralized mining in centralized pools," Review of Financial Studies, 34 (3), 1191–1235.

Cong, Lin William, George Andrew Karolyi, Ke Tang, and Weiyi Zhao (2021b): "Value Premium, Network Adoption, and Factor Pricing of Crypto Assets," Network Adoption, and Factor Pricing of Crypto Assets (December 2021).

Cong, Lin William, Wayne Landsman, Edward Maydew, and Daniel Rabetti (2023): "Tax-loss harvesting with cryptocurrencies," Journal of Accounting and Economics, 101607.

Cong, Lin William, Xi Li, Ke Tang, and Yang Yang (2021c): "Crypto wash trading," arXiv preprint arXiv:2108.10984.

Cong, Lin William, Ye Li, and Neng Wang (2021d): "Token-based platform finance," Journal of Financial Economics.

——— (2021e): "Tokenomics: Dynamic adoption and valuation," Review of Financial Studies, 34 (3), 1105–1155.

Cong, Lin William and Simon Mayer (2021): "The Coming Battle of Digital Currencies," Available at SSRN 3992815.

Cong, Lin William, Ke Tang, Yanxin Wang, and Xi Zhao (2022): "Inclusion and Democratization Through Web3 and DeFi? Initial Evidence from the Ethereum Ecosystem," Working Paper.

Daniel, Kent, Robert Hodrick, and Zhongjin Lu (2017): "The Carry Trade: Risks and Drawdowns," Critical Finance Review, 6 (2), 211–262.

Duffie, Darrell (2010): Dynamic asset pricing theory, Princeton University Press.

Easley, David, Maureen O'Hara, and Soumya Basu (2019): "From mining to markets: The evolution of bitcoin transaction fees," Journal of Financial Economics, 134 (1), 91–109.

Fama, Eugene F (1984): "Forward and spot exchange rates," Journal of monetary economics, 14 (3), 319–338.

Fama, Eugene F and Kenneth R French (1998): "Value versus growth: The international evidence," The journal of finance, 53 (6), 1975–1999.

Fan, Zhenzhen, Feng Jiao, Lei Lu, and Xin Tong (2024): "The Risk and Return of Cryptocurrency Carry Trade," Working Paper.

Fanti, Giulia, Leonid Kogan, and Pramod Viswanath (2021): "Economics of proof-of-stake payment systems," in Working paper.

Feenstra, Robert C (1986): "Functional equivalence between liquidity costs and the utility of money," Journal of Monetary Economics, 17 (2), 271–291.

Franz, Friedrich-Carl and Alexander Valentin (2020): "Crypto Covered Interest Parity Deviations," Available at SSRN 3702212.

Gabaix, Xavier (2009): "Power laws in economics and finance," Annu. Rev. Econ., 1 (1), 255–294.

Gabaix, Xavier and Matteo Maggiori (2015): "International liquidity and exchange rate dy-

namics," The Quarterly Journal of Economics, 130 (3), 1369–1420.

Gans, Joshua S, Hanna Halaburda, et al. (2015): "Some economics of private digital currency," Economic analysis of the digital economy, 257–276.

Griffin, John M, Xiuqing Ji, and J Spencer Martin (2003): "Momentum investing and business cycle risk: Evidence from pole to pole," The Journal of finance, 58 (6), 2515–2547.

Griffin, John M and Amin Shams (2020): "Is Bitcoin really untethered?" Journal of Finance, 75 (4), 1913–1964.

Harvey, Campbell R, Ashwin Ramachandran, and Joey Santoro (2021): DeFi and the Future of Finance, John Wiley & Sons.

Hinzen, Franz J, Kose John, and Fahad Saleh (2019): "Proof-of-work's limited adoption problem," NYU Stern School of Business.

Hou, Kewei, G Andrew Karolyi, and Bong-Chan Kho (2011): "What factors drive global stock returns?" The Review of Financial Studies, 24 (8), 2527–2574.

Howell, Sabrina T, Marina Niessner, and David Yermack (2020): "Initial coin offerings: Financing growth with cryptocurrency token sales," Review of Financial Studies, 33 (9), 3925–3974.

Huberman, Gur, Jacob D Leshno, and Ciamac Moallemi (2021): "Monopoly without a monopolist: An economic analysis of the bitcoin payment system," Review of Economic Studies, 88 (6), 3011–3040.

Ilmanen, Antti (1995): "Time-varying expected returns in international bond markets," The Journal of Finance, 50 (2), 481–506.

Jermann, Urban J (2023): "A Macro Finance Model for Proof-of-Stake Ethereum," Available at SSRN 4335835.

Jiang, Zhengyang, Arvind Krishnamurthy, and Hanno Lustig (2021): "Foreign safe asset demand and the dollar exchange rate," The Journal of Finance, 76 (3), 1049–1089.

John, Kose, Thomas J Rivera, and Fahad Saleh (2020): "Economic implications of scaling blockchains: Why the consensus protocol matters," Available at SSRN 3750467.

——— (2022): "Equilibrium staking levels in a proof-of-stake blockchain," Available at SSRN 3965599.

Kogan, Leonid, Giulia Fanti, and Pramod Viswanath (2021): "Economics of proof-of-stake payment systems," .

Koijen, Ralph SJ, Tobias J Moskowitz, Lasse Heje Pedersen, and Evert B Vrugt (2018): "Carry," Journal of Financial Economics, 127 (2), 197–225.

Lasry, Jean-Michel and Pierre-Louis Lions (2007): "Mean field games," Japanese journal of mathematics, 2 (1), 229–260.

Lehar, Alfred and Christine A Parlour (2020): "Miner collusion and the bitcoin protocol," Available at SSRN 3559894.

Li, Tao, Donghwa Shin, and Baolian Wang (2021): "Cryptocurrency pump-and-dump schemes," Available at SSRN 3267041.

Li, Tao, Chuyi Sun, Donghwa Shin, and Baolian Wang (2022): "The Dark Side of Decentralized Finance," Working Paper.

Liu, Yukun, Aleh Tsyvinski, and Xi Wu (2019): "Common risk factors in cryptocurrency," Tech. rep., National Bureau of Economic Research.

Lustig, Hanno, Nikolai Roussanov, and Adrien Verdelhan (2011): "Common risk factors in

currency markets," The Review of Financial Studies, 24 (11), 3731–3777.

——— (2014): "Countercyclical currency risk premia," Journal of Financial Economics, 111 (3), 527–553.

Lustig, Hanno, Andreas Stathopoulos, and Adrien Verdelhan (2019): "The term structure of currency carry trade risk premia," American Economic Review, 109 (12), 4142–77.

Lyandres, Evgeny, Berardino Palazzo, and Daniel Rabetti (2019): "Do tokens behave like securities? An anatomy of initial coin offerings," SSRN Electronic Journal.

Menkhoff, Lukas, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf (2012): "Carry trades and global foreign exchange volatility," The Journal of Finance, 67 (2), 681–718.

Merton, Robert C (1971): "Optimum consumption and portfolio rules in a continuous-time model," Journal of Economic Theory, 3 (4), 373–413.

Pardoux, Etienne and Aurel Râșcanu (2014): Stochastic differential equations, Backward SDEs, Partial differential equations, vol. 69, Springer.

Park, Andreas (2021): "The conceptual flaws of constant product automated market making," Available at SSRN 3805750.

Petersen, Mitchell A (2008): "Estimating standard errors in finance panel data sets: Comparing approaches," The Review of financial studies, 22 (1), 435–480.

Prat, Julien, Vincent Danos, and Stefania Marcassa (2019): "Fundamental pricing of utility tokens," .

Saleh, Fahad (2021): "Blockchain without waste: Proof-of-stake," Review of Financial Studies, 34 (3), 1156–1190.

Schmeling, Maik, Andreas Schrimpf, and Karamfil Todorov (2023): "Crypto carry," Available at SSRN 4268371.

Tang, Ke and Wei Xiong (2012): "Index investment and the financialization of commodities," Financial Analysts Journal, 68 (6), 54–74.

Tang, Ke and Haoxiang Zhu (2016): "Commodities as collateral," The Review of Financial Studies, 29 (8), 2110–2160.

Valchev, Rosen (2020): "Bond convenience yields and exchange rate dynamics," American Economic Journal: Macroeconomics, 12 (2), 124–66.

Verdelhan, Adrien (2010): "A habit-based explanation of the exchange rate risk premium," The Journal of Finance, 65 (1), 123–146.

Wold, Herman OA and Peter Whittle (1957): "A model explaining the Pareto distribution of wealth," Econometrica (Pre-1986, Journal of the Econometric Society), 591–595.

Online Appendices for "The Tokenomics of Staking"

OA1 Institutional Background of Staking

OA1.1 Staking Mechanisms

Blockchain-based staking in general involves two broad categories of activities: those related to pan-PoS consensus protocols and those in higher layer DeFi applications.[49] Even on non-blockchain-based or centralized platforms, various programs that involve escrows or crowd funds can be analyzed as a form of business layer staking through the lens of our framework. Fundamentally, a blockchain functions to generate a relatively decentralized consensus record of system states to facilitate economic interactions such as value or information exchanges (e.g., Cong and He, 2019). PoS protocols have gained popularity and momentum with major market players such as Ethereum adopting them. Under PoS, agents stake native tokens to compete for the opportunity to record transactions, execute smart contracts, append blocks, etc., to earn block rewards and fees. Meanwhile, various staking programs have become popular means to encourage desirable behavior in higher layer applications, escrowing a balance of tokens under custody in a smart contract, or deploy them to enable network economic functionalities.

Consensus generation in PoS. Fundamentally, blockchain functions to generate a relatively decentralized consensus to enable economic interactions such as value or information exchanges (e.g., Cong and He, 2019). Permissionless blockchains have historically relied on variants of the PoW protocol. Because of scalability and environmental issues of PoW (Cong et al., 2021a; John et al., 2020), PoS protocols have gained popularity and momentum for both permissioned and permissionless blockchains, with major market players adopting and incumbents such as Ethereum contemplating a conversion (Irresberger et al., 2021).

Under PoS, agents who stake native tokens have the opportunity to append blocks and earn block rewards and fees as compensation. The more one stakes, the more likely one is to be selected and compensated for their participation (Saleh, 2021, contains more details). Note that holding a token does not necessarily mean participating in staking. In practice, agents incur negligible physical costs (as opposed to the high entry cost of PoW mining or directly maintaining a node in PoS). Our study includes all protocols using pan-PoS protocols, such as Proof-of-Credit (POC) used in Nuls, which are variants of the above mechanisms. Online Appendix OA1.2 provides more background information and examples in practice.

Staking in DeFi. Staking programs are a popular means for incentivizing desirable behavior and guarding against misbehavior in DeFi applications. It escrows a balance of tokens under custody in a smart contract and stakers receive rewards similar to interest payments from their tokens staked (Harvey et al., 2021). Synthetix is an example of an open-source DeFi protocol with staking where users can create and trade derivative tokens and gain exposure to assets like gold, bitcoin, and euros without having to actually own them. These derivative assets are collateralized by the platform tokens (SNX) which, when locked in the contracts, enables their issuance. In return, SNX stakers earn rewards from both newly issued tokens and small fees transactions generate.[50] Another

[49] The two are not mutually exclusive. Solana, for example, uses both PoS and DeFi staking. The classification we use follows mainstream cryptocurrency data aggregators such as CoinMarketCap.

[50] DeFi staking may involve multiple tokens. For example, in MakerDAO, the profits generated from DAI can be viewed as a yield on ETH staking, and our framework can be used to understand the price impact on ETH.

salient form of staking is yield farming, which allows investors to earn yield by temporarily locking tokens in a decentralized application (dApps). Yield farming often entails shorter lock-up (some allows withdrawal at any time), uncertain yields, and higher risks. (see, Augustin et al., 2022, for more institutional details).

Without getting bogged down with specific eligibility requirements and operational differences across various DeFi protocols and smart contracts, DeFi staking can be characterized as simply earning rewards by collateralizing the tokens for some functionalities in the network.

Reward determination and slashing. In most staking programs, the total rewards used to incentivize staking or its determination mechanism are pre-specified and announced. In PoS, the blockchain branch is randomly selected from the whole staking pool. Then the staking reward is randomly distributed to stakeholders based on the number of staked coins they hold as a probability weight.[51] Similarly, on DeFi platforms, stakers share the rewards from transaction fees or predetermined emissions (minting of new tokens).

Staking reward rate can be naturally compared to interest rate or yield of other financial assets. However, unlike deposit rates set by the banks, staking reward rate is jointly determined by the announced staking reward and the aggregate tokens staked. Online Appendix OA1.2 details the staking programs for the tokens in our sample. Stakers face the risk of losing the staked tokens due to possible security attacks, illegal verification, and storage failures. To discourage the misbehavior of the validator, most projects also propose a punishment mechanism known as *slashing*. A pre-defined percentage of a validator's tokens are lost when it does not behave consistently or as expected on the network (e.g., downtime and double signing).

Market and information. In PoS, validators compete in the amount of staking to earn rewards. To incentivize more delegates, they develop a reward distribution plan at the node level. Potential delegates can freely choose among these nodes or delegate through some intermediaries. Therefore, nodes engage in price competition for delegated stakes. For DeFi platforms, staking reward rates are typically equal for participants, but some white-listed groups may have priority in staking. Most stakable tokens are launched on mainstream cryptocurrency exchanges. Investors can easily invest in these staking projects and trade these tokens with cryptocurrency assets such as Bitcoin and Ethereum.

Information on staking programs, including participation rules, reward distribution plan, total staked value (or total value locked, TVL, which includes non-native tokens), and even information of all the validators, are open and can be easily obtained on official websites of projects. Third-party websites also specialize in collecting real-time information on staking projects, e.g., *Stakingrewards.com*. In particular, the *staking ratio*, which captures the total number of tokens staked as a fraction of the total number of tokens, is typically public knowledge.

OA1.2 Forms of Staking in Our Sample

PoS staking. Under PoS, agents who stake native tokens have opportunities to append blocks and earn block rewards and fees as compensation. There are mainly two ways to participate. The first is to *run a validator node, staking pool, or masternode* by holding native tokens and incurring the costs including hardware costs and time spent on maintenance. The more one stakes, the more likely one is to be selected and compensated

[51] For example, if an investor stakes 10 coins while the aggregate staked amount of this branch is 100, then the investor has a 10% probability of appending to the branch and receiving the staking reward.

for their participation(Saleh, 2021, contains more details). Note that holding a token does not necessarily mean participating in staking. The second way is through *delegation*. Agents only need to delegate their tokens to an existing node or a pool to receive a reward earned by the node/pool. This route is flexible and friendly for players with fewer tokens and allows them to share risk (Cong et al., 2021a). In practice, agents incur negligible physical costs (as opposed to the high entry cost of PoW mining or directly maintaining a node in PoS).

Solana is a concrete example of pan-PoS staking.[52] Solana is an open-source project that implements a new, high performance, permissionless blockchain. It enables transactions to be ordered as they enter the network, rather than by block, which makes Solana one of the fastest blockchains in the world and the rapidly growing ecosystem in crypto, with thousands of projects spanning DeFi, NFTs, Web 3.0 and more. Solana uses Proof-of-Stake (PoS) as its consensus mechanism. The performance is improved by its innovative protocol, Proof-of-History (PoH). Solana's Proof-of-Stake is designed to quickly confirm the current sequence of transactions produced by the PoH generator, vote and select the next PoH generator, and punish misbehaving validators. A block in the context of Solana is simply the term used to describe the sequence of entries that validators vote on to achieve confirmation. *Validators* within Solana's PoS consensus model are the entities responsible for confirming if these entries are valid. *SOL* is the name of Solana's native token, which can be passed to nodes in a Solana cluster in exchange for running an on-chain program or validating its output. *Stakers* delegate SOL to validators to help increase these validators' voting weight. Such action indicates a degree of trust in the validators. Stakers delegate to ensure validators cast honest votes and hence ensure the security of the network. The more stake delegated to a validator, the more often this validator is chosen to write new transactions to the ledger, and then the more *rewards* the validator and its delegators earn.

Staking DeFi native tokens. Incentivizing desirable behavior and guarding against misbehavior are crucial in DeFi applications. To this end, staking programs are popular and important in practice, which applies to a balance of tokens under custody in a smart contract. Users on DeFi platforms receive staking rewards as a form of interest payment from their token balance staked (Harvey et al., 2021).

In practice, DeFi staking may involve different lock-up periods and multiple tokens.[53] The risks of being slashed and losing the staked tokens are also different. Without getting bogged down with specific threshold requirements and operational differences across various DeFi protocols and smart contracts, DeFi staking can be characterized as simply earning rewards by collateralizing the tokens for some functionalities in the network. From the stakers' perspective, staking shares the spirit of certificates of deposit or risky illiquid investments.

Transaction gas fees and rewards. We summarize the gas fee foundations of Ethereum (EIP 1559) as an example. The contents are primarily extracted from the official webpages of Ethereum and *blocknative*.[54]

Gas refers to the unit that measures the amount of computational effort required to execute specific operations on the Ethereum network. Since each Ethereum transaction requires computational resources to execute, each transaction requires a fee. Gas thus refers to the fee required to execute a transaction on Ethereum.

[52]For references, see https://docs.solana.com, *blockdaemon.com*, and https://blockdaemon.com/platform/validator-node/how-solana-staking-works/.

[53]MakerDAO is a good example. The profits generated from DAI can be viewed as a yield on ETH staking, and our framework can be used to understand the price impact on ETH.

[54]Please see https://ethereum.org/en/developers/docs/gas/, and https://www.blocknative.com/blog/eip-1559-fees.

Every block has a *base fee* which acts as a reserve price. To be eligible for inclusion in a block the offered price per gas must at least equal the base fee. The base fee is calculated independently of the current block and is instead determined by the blocks before it - making transaction fees more predictable for users. When the block is mined this base fee is "burned," removing it from circulation. The base fee is calculated by a formula that compares the size of the previous block (the amount of gas used for all the transactions) with the target size. The base fee will increase by a maximum of 12.5% per block if the target block size is exceeded. This exponential growth makes it economically non-viable for block size to remain high indefinitely.

In origin, miners would receive the total gas fee from any transaction included in a block. With the new base fee getting burned, the London Upgrade introduced a *priority fee* (tip) to incentivize miners to include a transaction in the block. Without tips, miners would find it economically viable to mine empty blocks, as they would receive the same block reward. Under normal conditions, a small tip provides miners a minimal incentive to include a transaction. For transactions that need to get preferentially executed ahead of other transactions in the same block, a higher tip will be necessary to attempt to outbid competing transactions.

Staking rewards are the combination of gas fees and the emission of new tokens. Together with the gas fee policies summarized above, as well as the mechanism of rewards from emission described in the main text, the main take-away is that the reward distribution design is widely under the platform's control, whereas the crowd decisions and interactions provide influence under the designed structure. In practice, the staking rewards may also involve the phenomenon of multilevel distribution, as some agents can delegate tokens to larger nodes. The holders of the larger nodes thus need to have a process for deciding on the distribution of benefits. Therefore, there may exist multiple staking participation methods for one token. In relevant empirical tests, we always choose the participation method with the lowest capital threshold and risk, such as delegating, voting, etc.

Examples in practice. We summarize representative staking programs involving tokens in our sample. Most information is accessed from *StakingRewards*. There is also information from official websites of corresponding tokens. Many tokens have similar mechanisms, thus we do not repeat the description. These descriptions are excerpted in 2022. There may be changes over time in the specific mechanisms of some programs, whereas these descriptions apply to the time intervals covered by the data in our paper.

- The individual AION rewards depends on the Block Reward, Block Time, Daily Network Rewards and Total Staked. Every block one validator is randomly selected to create a block, whereas 1 staked or delegated token counts as one "lottery ticket." The selected validator has the right to create a new block and broadcast them to the network. The validator then receives the 50% of the block reward and the fees of all transactions (network rewards) successfully included in this block, whereas the PoW Miner receives the other 50%.

- Rewards in the form of algos are granted to Algorand users for a variety of purposes. Initially, for every block that is minted, every user in Algorand receives an amount of rewards proportional to their stake in order to establish a large user base and distribute stake among many parties. As the network evolves, the Algorand Foundation will introduce additional rewards in order to promote behavior that strengthens the network, such as running nodes and proposing blocks.

- The individual BitBay rewards depends on the Block Reward, Block Time, Daily Network Rewards and

Total Staked. Every block is randomly selected whereas 1 staked coin counts as one "lottery ticket." The selected staker has the right to create a new block and broadcast it to the network. He then receives the block reward and the fees of all transactions successfully included in this block.

- Dash blockchain consensus is achieved via Proof of Work + Masternodes. Investors can leverage their crypto via operating masternodes. Miners are rewarded for securing the blockchain and masternodes are rewarded for validating, storing and serving the blockchain to users.
- Eos has a fixed 5% annual inflation. 4% goes to a savings fund, which might distribute the funds to the community later on. 1% goes to Block producers and Standby Block Producers. Out of the 1% that are given to block producers, only 0.25% will go to the actual 21 producers of the blocks. The other 0.75% will be shared among all block producers and standby block producers based on how many votes they receive and with a minimum of 100 EOS/day.
- The individual reward of staking fantom depends on the Total Staked ratio. Transactions are packaged into event blocks. In order for event blocks to achieve finality, event blocks are passed between validator nodes that represent at least 2/3rds of the total validating power of the network. A validator's total validating power is primarily determined by the number of tokens staked and delegated to it. A validator earns rewards each epoch for each event block signed according to it's validating power. By delegating, investors can increase the share of their validator proportionally to the balance of their account. They will receive rewards accordingly and share them with investors after taking the commission.
- The effective yield for staking IDEX depends on the actual Trading Volume on IDEX Market. The higher the trading volume on IDEX, the higher are the actual rewards. The second metric to watch is the total amount of AURA currently staking. Fewer tokens on stake result in higher rewards.
- Every livepeer (LPT) token holder has the right to delegate their tokens to an Orchestrator node for the right to receive both inflationary rewards in LPT and fees denominated in ETH from work completed by that node.
- The individual LTO rewards depends on the Network Rewards (Transaction Fees spent on the Network) and the Total Staked. Every block one staking node operator is randomly selected to create a new block, whereas 1 staked token counts as one "lottery ticket." The staker receives the fees of all transactions successfully included in this block. Staking Node Operators share the rewards with their delegators after deducting a commission.
- NEM blockchain consensus is achieved via Proof of Importance. Investors can leverage their crypto via harvesting. To harvest NEM coins it is recommended to run the official NEM Core wallet with an entire copy of the blockchain on the stakers' computer or a Virtual Private Server (VPS). The individual NEM harvesting rewards depends on the Daily Network Rewards and Total Staked. For every block, the staker is randomly selected whereas 1 staked coin counts as one "lottery ticket." The selected staker has the right

to create a new block and broadcast it to the network. The staker then receives the fees of all transactions successfully included in this block.

- Everyone who holds NEO will automatically be rewarded by GAS. GAS is produced with each new block. In the first year, each new block generates 8 GAS, and then decreases every year until each block generates 1 GAS. This generation mechanism will be maintained until the total amount of GAS reaches 100 million and no new GAS will be generated.

- Nuls blockchain consensus is achieved via Proof of Stake + Masternodes. Investors can leverage their crypto via staking. The amount earned is variable based on the current blockchain metrics like the amount of stakers (Total Staked ratio). Investors can stake Nuls into a project's nodes and earn their token as a reward, while the project earns Nuls as a reward. Some projects offer to stake with just 5 Nuls as the minimum.

- Delegators in Polkadot are called Nominators. Anyone can nominate up to 16 validators, who share rewards if they are elected into the active validators set. The process is a single-click operation inside the wallet. The current reward rate for validators is determined by the current Total Staked ratio. The less DOT is being staked, the higher are the rewards.

- Qtum blockchain consensus is achieved via Proof of Stake 3.0. The individual reward depends on the Block Reward, Block Time, Daily Network Rewards and Total Staked. Every block is randomly selected whereas 1 staked coin counts as one "lottery ticket." The selected staker has the right to create a new block and broadcast it to the network. The staker then receives the block reward and the fees of all transactions successfully included in this block.

- Synthetix Network Token blockchain consensus is achieved via the Ethereum Blockchain. Investors can leverage their crypto via staking. SNX holders can lock their SNX as collateral to stake the system. Synths are minted into the market against the value of the locked SNX, where they can be used for a variety of purposes including trading and remittance. All Synth trades on Synthetix Exchange generate fees that are distributed to SNX holders, rewarding them for staking the system.

- Tezos blockchain consensus is achieved via Liquid Proof of Stake. Investors can leverage their crypto via baking or delegating. There are a number of tokens that use a similar mechanism, including iotex, irisnet, etc.

- Tron reward depends on the Block Rewards, Endorsement Rewards, Block Time, Daily Network Rewards and Total Staked. Every block is randomly selected to bake a block and 32 stakers are selected to endorse a block, whereas 1 staked coin counts as one "lottery ticket." The selected stakers have the right to create or endorse new block and broadcast them network. The Baker then receives the block reward and the fees of all transactions successfully included in this block. The Endorsers receive the endorsement rewards.

- Wanchain blockchain consensus is achieved via Galaxy Proof-of-Stake. The individual WAN rewards depends on the Foundation Rewards, Daily Network Rewards and Total Staked. At the beginning of each

protocol cycle (epoch), two groups, the RNP (Random Number Proposer) group and the EL (Epoch Leader) group, are selected from all validators. 1 staked or delegated token counts as one "lottery ticket" to be selected. The two groups equally share the Foundation Rewards and Transaction Fees (Network Rewards). The Foundation Rewards consists of 10% of the outstanding Wanchain Token Supply and are decreasing by 13.6% each year, whereas the Network Rewards are expected to rise alongside wider network usage.

OA2 Notes on Numerical Solutions

We document the numerical approaches in Section 4.4. In brief, we connect Bilal (2023)'s idea of analytical perturbation and the approximation using instantaneous arriving shocks, and then derive a system that solvable by the finite difference (FD) method.

In the master equation (25), the agents do not respond to any specific shocks, but only "internalize" them by the penalty term as well as the volatility risk of penalty. To better comprehend the trajectory of the economy after specific shocks, we adopt the underlying simulation idea. Suppose once an agent decides on an allocation, she cannot adjust her portfolio choice within a short time interval $\mathrm{d}t$, regardless of the magnitude of the shock that realizes during this brief period. Thus, z is presumed when solving optimization problems and enters the wealth dynamic as well as the penalty function as an external parameter. In simpler terms, the uncertainty over the entire future time horizon is dissected into two components as $\mathrm{d}t \to 0$: the part that remains unrealized, which operates as a risk affecting agents' retrospective evaluation, and the part that corresponds to the realized shock z. This realized shock z is incorporated into the optimization process as an additional external parameter influencing wealth dynamics and the penalty structure. This inclusion is essential since, given a specific z, agents can only hypothesize its value when solving optimization problems.

Then we show how z enters the value function J. Take $\mathrm{d}t \to 0$ and the weighted traversal of any possible z into account, the resulting value function is proved to be an appropriate approximation of that in the initial master equation. Recall the master equation (25) contains the derivatives w.r.t. m. Though it is conceptually not difficult to calculate the approximation of the second-order derivative, it generates a large bracket of terms but bring relatively unclear general insights. What we clearly know is that it comes from the Brownian uncertainty on the distribution, i.e., $\frac{\partial}{\partial w}[g(w)m(w)\mathrm{d}Z_t]$. Alternatively, we first approximate the master equation from the economic perspective, i.e., consider the following scenario. At time t and near the steady state, each agent solves the optimization problem given the current platform productivity A and distribution m. The difference from the background of the above master equation is that once she determines the allocation, she cannot change her portfolio choice in a short time $\mathrm{d}t$, no matter what degree the shock in this short period realizes. In other words, the uncertainty in the whole future time dimension is divided into two parts as $\mathrm{d}t \to 0$: The unrealized part performs as a risk that affects agents' backward induction. The realized shock z enters the optimization directly as an extra exogenous argument, since for any given z, agents can only presuppose its value when solving optimization problems. An implicit requirement that allows this approach to be a proper approximation is that the generator of uncertainty is memoryless. Also, this approach shares a similar idea with Bertucci (2021), which rigorously proves that under the situation as our model introduces, i.e., all the agents are pushed by the same Brownian motion, it is a good approximation by introducing a transformation that is going to affect all players in the game at random times.

Then the law of motion is

$$\mathrm{d}m_t/\mathrm{d}t = L^*(w)[m] - \frac{\partial}{\partial w} g(w)m(w)z \equiv L^*(w)[m] + zG^*(w)[m], \tag{OA.1}$$

where $G^*(w)[m] = -\frac{\partial}{\partial w}[g(w)m(w)]$. In addition, by the separate form of the value function J, we only focus on the part that does not involve A and denote the resulting part of the value function as the new J for simplicity. Then substituting into the master equation above and calculating the penalty term, we obtain

$$\begin{aligned} \phi J(w,m,z) = \mathcal{U}(w,m,z) + L(w)[J] + \int \frac{\delta J}{\delta m}(w,w',m)L^*(w')[m]\mathrm{d}w' \\ + \frac{\partial}{\partial w}\left[g(w)\int \frac{\delta J}{\delta m}(w,w',m)\frac{\partial g(w')m(w')}{\partial w'}\mathrm{d}w'\right] + \mathcal{Z}(z)[J], \end{aligned} \tag{OA.2}$$

where $\mathcal{Z}(z)[J] = \frac{1}{2}\frac{\partial^2 J}{\partial z^2}$. This term is obtained from the fact that agents also take the generation of such a shock into account. That is, the agent not only presupposes the case of such z, but also takes into account the possibility of such a presupposition. Note that the direct impact of z, i.e. the corresponding term in (OA.1) is offset by the penalty vz, while the penalty generates the additional term $\frac{\partial}{\partial w}g(w)v(w)$. Also, the precise form of $\mathcal{Z}(z)[J]$ is determined by the generator of Z_t. Since we have accounted for σ in $g(w,m)$, the rest part is a standard Brownian motion.

So far, we have obtained a version of value function that contains a presumed shock z. By combining all the possible shocks, we simulate the original value function. Thus, the only thing to do is to solve $J(w,m,z)$ in (OA.2). On solving $J(w,m,z)$, we follow Bilal (2023) to adopt a perturbation method near the steady state. Let $v(w,w') = \frac{\delta J}{\delta m}(w,w',m^{SS})$ ("SS" refers to the steady state), which can be also understood as the "deterministic" impulse value function. Apply the similar method of analytical perturbation in Bilal (2023), we set a small distributional shock h as well as a relatively small scale to the shock $\epsilon = \tilde{Z}/z$, $\epsilon h = m - m^{SS}$, then the first order solution to the master equation (OA.2) reads

$$J(w,m,z) = J^{SS}(w) + \epsilon\left\{\int v(w,w')h(w')\mathrm{d}w' + \omega(w,z)\right\} + \mathcal{O}_w(||h||^2_{H^2}), \tag{OA.3}$$

where $v(w,w')$ refers to the deterministic part of the impulse value, whereas $\omega(w,z)$ is the stochastic part of the impulse. Bilal (2023) provides the complete explanation of the above process. The difference here is that both the deterministic and stochastic part come from the aggregated distributional shock, and correspond to the two parts of uncertainties as we mentioned above.[55] That is, the value function J is a linear functional of h to the first order.

The next step is then to solve the two impulse functions, $v(w,w')$ and $\omega(w,z)$. For $v(w,w')$: substituting into the master equation, we will obtain an equation w.r.t. the impulse value function, $v(w,w')$. Precisely, around the steady state, the associated vector that integrates against h must be zero. Then the approximate master equation is calculated by letting the first-order coefficients of h satisfy the equation that L.H.S. = R.H.S. We omit the expansions of those terms w.r.t. h but note the following observations: $\partial_J \mathcal{U}, \delta_m \mathcal{U}, \partial_J f, \partial_J g, \partial_m g = 0$, $\partial_w v(w,w')$

[55] Note that the first-order approximation is additive in the response to a distributional impulse h and the aggregate shock z. Any pairwise perturbation involving the both is second-order.

is second-order w.r.t. h, and the property of adjoint operators, $\int L(w)[a]b(w)\mathrm{d}w = \int a(w)L^*(w)[b]\mathrm{d}w$, and finally obtain

$$\begin{aligned}\phi v(w,w') = {} & \mathcal{L}_m(w,w')[J^{SS}] + \mathcal{L}(w)[v(\cdot,w')] + \mathcal{L}(w')[v(w,\cdot)] + \mathcal{G}(w')[v(w,\cdot)] \\ & + \int v(w,w'')\left\{\mathcal{L}_m^*(w'',w')[m^{SS}] + \mathcal{L}_y^*(w'',\mathcal{M}(w'',w',v))[m^{SS}]\right\}\mathrm{d}w'',\end{aligned} \tag{OA.4}$$

and similarly, $\omega(w,z)$ satisfies

$$\phi\omega(w,z) = z\left\{\mathcal{U}_z^{SS}(w,z) + \mathcal{L}_z(w)[J^{SS}]\right\} + \mathcal{L}(w)[\omega(\cdot,z)] + \mathcal{Z}(z)[\omega(w,\cdot)] + z\mathcal{G}_z^*(w'')[m^{SS}], \tag{OA.5}$$

where $\mathcal{M}$ is the distributional marginal propensity to control, implying how individual controls respond to a small distributional impulse, $\mathcal{G}^*(w')[m^{SS}] = -\frac{\partial}{\partial w'}\left(g(w')m^{SS}(w')\right)$, the subscripts of the operators are the Fréchet derivatives, e.g., $\mathcal{L}_m(w,w')[J^{SS}] = \frac{\delta}{\delta m}L(w,w')[J^{SS}]$.

Then, we only need to numerically solve (OA.4) and (OA.5). Note that they only contain the unknown matrices (in numerical processes), $v(w,w')$ and $\omega(w,z)$, and the terms at SS. Therefore, they are much easier to solve. Substituting into OA.2, the value function presuming z is solved. Then the original value function is obtained as the weighted aggregation of the cases presuming all possible z. However, the derivation of the impulse response for a particular sequence of shocks does not need to go this far — recall the resulting key changes, distributional disparity h. The perturbation method suggests that near the SS, the deterministic and stochastic impulse kernel, D and S, respectively satisfy

$$D(w,w') = \mathcal{L}_m^*(w,w')[m^{SS}] + \mathcal{L}_c^*(w_i,\mathcal{M}(w_i,w_j,v))[m^{SS}]; S(w,z) = z\mathcal{G}_z^*(w'')[m^{SS}]. \tag{OA.6}$$

Further, the impulse in the distribution h_t follows the SPDE,

$$\mathrm{d}h_t(w,z) = \{\mathcal{L}^*(w)[h_t] + \mathcal{D}(w)[h_t] + S(w,z_t)\}\,\mathrm{d}t. \tag{OA.7}$$

Then, given a series of realized shock $\{z\}_{t=1}^T$ and the initial h_0, the discretized distributional impulse is obtained by numerically solving the above equation.

Parameter choices. In numerical simulations, we calibrate the exogenous parameters based on existing literature and practice. For parameters related to platform productivity and transaction convenience, we adopt Cong et al. (2021b)'s approach. The productivity drift μ^A is set to be linear in θ, $\mu^A = \mu_0^A + a_{feed}(\Theta - 0.5)$, where $\mu_0^A = 2\%$ implying that when there is no feedback effect ($a_{feed} = 0$), the benchmark productivity drift is the same as Cong et al. (2021b)'s setting. a_{feed} is set to be 0 or 3% for comparative statics and is extensively discussed later. The Cobb-Douglas parameter α equals 0.3. The user type parameter $\beta = 0.8 < 1$, which captures the diminishing marginal transaction needs. $\hat{u}$ is set to be 10^{-4} and 10^{-5} for comparative statics when discussing difference tendencies to stake. Note that in the transaction convenience formula, A and U are homogeneous, we focus on exponential changes in $\hat{u}$ to make the comparison more visible. Since N_t is only required to be an increasing function of L_t and equals zero when $L_t = 0$, we use $N_t = L_t$ for simplicity. We also test the nonlinear forms, which brings up additional complex comparative statics beyond the focus of our study. The form of numéraire

convenience refers to Valchev (2020), and fits the assumptions in Section 2.1 and 3.2, i.e., Ψ is linear in A_t and numéraire holding with parameter $-1\%/e^{0.02}$ and -1%, respectively, where $e^{0.02}$ is used for adjusting the range of discussion on A_t to be comparable to the existing literature (e.g., Cong et al. (2021b)), and is not crucial for obtaining implications. We also consider the potential convenience loss if over-staking reduces liquidity. The parameters are adjusted by a multiplier $a_{loss}\Theta_t$, where $a_{loss} > 0$ meaning that numéraire convenience is relatively larger when the token has lower liquidity. a_{loss} has been examined differently while does not bring excess intuitions, and thus is not discussed in the main text. The staking-related parameters are calibrated to approximately fit corresponding indices in reality. Emission rate equals 1%, and $F_t = 1\%(1-e^{-\Theta_t})/\Theta_t$, which has similar properties with common fee designs as described in OA OA1.2, i.e., (i) prior fee decreases in aggregate scale of validators, and (ii) the fee rate has both lower and upper bounds. The demand shock has diffusion $\sigma^Z = 10\%$ so that the simulated price daily volatility is near 10% and close to our data sample. To be clear, because tokens in practice vary considerably in their design, we focus more on qualitative properties and have not pursued calibrating the data for a particular token over specific periods, although it may bring about additional quantitative results. In addition, agents' discount equals 1% in physical measure as commonly used in literature. The staking cost is set to 0.1% and is not crucial for analysis.

Supplementary discussion on the impulse responses. Figure OA.1 visualizes the impulse value functions, as well as the impulse response paths of unit shock. Panel A.(1)-(2) visualize the results of $v(w, w')$ and $\omega(w, z)$. The deterministic impulse value is widely negative, indicating the agents are competitive with each other. Further, the whales (agents with large w) are virtually unaffected by the retails, while their entry has huge impact to the retails. The only exception is the agents near the lowest threshold.[56] They are less affected by any other agents, since they have less onchain allocation. In subplot (2), for similar reasons, the retails have smaller exposure to the realized shock z.

As discussed, the numerical solving process allows us to simulate any given arrived series of shocks. On the other hand, standing at any period, the response is, to some extends, the combination of the continuous impacts of historical shocks. As such, we simulate the impulse sequence of a unit shock separately, which helps understand the relationship between continuously-arriving shocks and the continuously-moving states. Precisely, we assume a unit negative shock at $t = 0$, and zero shocks for any $t > 0$. As Figure OA.1 Panel (B) shows, the initial effect is a universal loss of money across all agents, leading to an increase in the share of retail agents and a decrease in the proportion of wealthier agents. Consequently, the overall staking ratio experiences a decline. As the response progresses, two driving forces come into play. Firstly, the affluent agents are compelled to further increase their wealth through both staking and transactions, thereby reinforcing their financial positions. Secondly, the positive adjustment of staking reward rates contributes to the staking ratio's response dynamics. Interestingly, there is an over-response to the staking ratio. In addition to the complex aggregation among the changing crowd, it is also explained by the platform development: Due to the decrease in staking, the progression of the platform's productivity A_t decelerates. As time advances, the value of A_t remains lower than it would be without the occurrence of such a shock. In comparison to the steady state, agents exhibit relatively higher preference in

[56] Recall (27), the distribution is similar with Pareto distribution, in which the minimal w is also needed. From the economic perspective, agents with too little wealth will be blocked by the threshold of transaction, φ, and have no onchain activities.

staking. This propensity persists until the developmental disparity is filled.

The numerical approach can also be used for simulating the direct impulse on the distribution. It could, for example, correspond to the cases where the economy suddenly changes by an unintended entry of external groups, especially the whales. This is also within the scope of our framework. That is, we suppose there is an unanticipated distributional shock h and examine its influence. As Figure OA.1 Panel (C) shows, the entry of a crowd of whales initially definitely induces a large rise in the proportion of affluent agents and an accompanying decline in the share of retail participants. This shift triggers a substantial increase in the staking ratio. However, the subsequent evolution of the wealth distribution exhibits a fatter tail, which aligns with Figure 3. This suggests that the primary benefit stemming from the entry of large funds lies in its potential to stimulate platform growth. Nevertheless, the competitive interaction among agents implies that such an entry invariably exerts impact on the on-chain inequality.

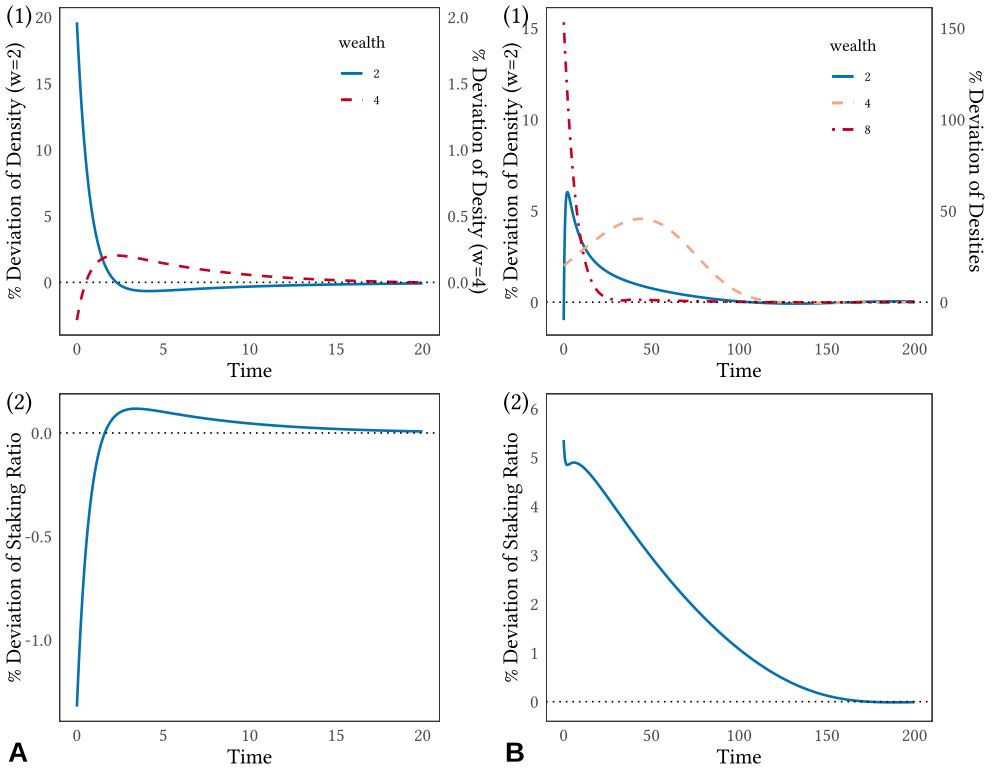


Figure OA.1: Impulse values and response functions: Supplementary tests.
Panel A.(1)-(2) show the deterministic and stochastic impulse value functions, $v(w, w')$ and $\omega(w, z)$, respectively. Panel B shows the impulse responses of a unit negative shock at $t = 0$, in which subplot (1) shows the response, i.e, the deviations comparing with the steady state of density $m(w)$ of different w, and subplot (2) shows the response of the staking ratio. Time unit is roughly calibrated to one-day. Similarly, Panel C shows the experiment when there is a crowd of whales near $w = 5$ enters the economy.

OA3 Background and Extended Discussions

OA3.1 Mean Field Game and the Master Equation

Mean field games (MFGs), introduced in the pioneering works of Lasry and Lions (2007), offer a powerful framework for analyzing strategic interactions in large populations when each individual agent has only a small impact on the behavior of other players. MFG supposes that the rational agents are indistinguishable and individ-

ually have a negligible influence on the game, and that each individual strategy is influenced by some averages of quantities depending on the states of the other agents. A very nice introduction to the theory of MFGs is supplied in the notes of Cardaliaguet (2010), including theoretical results on the existence and uniqueness of classical solutions, and also discussions on weak solutions.

MFG has a wide range of potential applications in economics. In macroeconomics, it has been applied to the studies that connects represent agent's optimization and the dynamics of macro interests, such as the income distribution (Achdou et al., 2022). It also allows heterogeneous settings, such as the heterogeneous-agent model (Krusell and Smith, 1998). Literatures also attempt to model financial problems using MFG. Brunnermeier and Sannikov (2016) compare the historical evolutions of macroeconomic and finance models, arguing that properly framed, the analysis of continuous time stochastic models should provide a unifying thread for these subfields of economics which so far, developed in parallel.[57] To this end, the authors introduce models of the economy comprising households maximizing consumption like in classical macroeconomic growth models, as well as investors trading in financial markets. Some applications have been discussed, including trade crowding (Cardaliaguet and Lehalle, 2018) and crypto mining (Li et al., 2019).

In many interesting situations in financial studies, it is important to allow for systematic risks (or systematic/common shocks). Such applications call for a more general framework in theory. Fortunately, the MFG system can be written in the most general case in terms of a so-called "Master Equation" (Cardaliaguet et al., 2019). The master equation is an equation on the space of measures; i.e. it is an equation that is set in infinite-dimensional space. The logic why the problem with aggregate uncertainty becomes infinite-dimensional is that the cross-sectional distribution across agents becomes a state variable in agents' dynamic programming problems and that distribution is an infinite-dimensional object.[58] The master equation is first introduced by Lions (2011). The most related research advances include the proof of the existence and uniqueness of a classical solution to the master equation (Cardaliaguet et al., 2019), monotonous solutions in several specific cases of common shocks (Bertucci, 2021), the case without idiosyncratic but with Brownian-type common shock (Cardaliaguet and Souganidis, 2020), and the corresponding weak solutions (Cardaliaguet and Souganidis, 2021).

The relationship between MFGs and the master equation. There are two lines for understanding the relationship between the two: The first one is in line with the development of the theory, whereas the second is relatively smoother from the perspective of applications in economics.

Consider the first logic. A MFG system is enough to handling situations where heterogeneous players solve optimal decisions even with idiosyncratic shocks and the crowd evolves dynamically. Although they can also be rearranged in the master equation form, the couple of PDEs shows clearly the respective dynamics of J and m, and fits mathematical tools for forward-backward PDEs. Therefore, the MFG system is usually described by the HJB equation together with the law of motion (FP equation). However, when the shock is not idiosyncratic but universal, it directly adds uncertainty to the distribution. Thus the system of PDEs both becomes stochastic as shown in (24). Even worse, the appearance of the extra "penalty" term makes the respective dynamics unclear in the separate PDEs. Therefore, the convenience of the coupled-PDEs form no longer exists. To this end, the

[57] This is also reviewed by Carmona (2020).

[58] This is also reviewed in the online Appendix of Achdou et al. (2022).

master equation is presented as a more generalized representation that applicable to aggregate shocks, while at the cost of the value equation needing to be defined over infinite dimensions.

The second logic is more in line with economic research. Many theoretical frameworks consider the representative agent who solves optimal choices under system states with shocks. The distribution of the crowd m_t could be one of the system states and enters the value function as an argument. From the formal perspective, we still totally differentiate the value function to obtain the "HJB equation." The thing is that m_t is a function itself and $\mathrm{d}m_t$ is actually the law of motion. Ignoring the technical details for the moment, we notice that the total differential involves the law of motion, and the resulting new "HJB equation" is essentially (at least formally) the master equation.

The two understanding logics also make slight differences in the derivations of the master equation. The total differential approach seems natural while faces with more rigor considerations, whilst the former aims to rearrange the PDEs and find the solution to the MFG system. The master equation is finally the equation that characterize the solution.

OA3.2 Derivation of the Master Equation

Relevant definitions and requirements. In the main text, we mentioned the notations, $\frac{\delta J}{\delta m}$ and $\frac{\delta^2 J}{\delta m^2}$, i.e., the first and second Fréchet derivatives applied in infinite dimension w.r.t. the density m. They appear to have extra arguments, i.e.,

$$\frac{\delta J}{\delta m} = \frac{\delta J}{\delta m}(w, w', m), \quad \frac{\delta^2 J}{\delta m^2} = \frac{\delta^2 J}{\delta m^2}(w, w', w'', m). \tag{OA.8}$$

One can roughly compare them with Jacobian and Hessian in the finite-dimensional case respectively. To gain rigorous requirements of the derivatives in a more generalized way, we introduce the Monge-Kantorovich distance first. The set $\mathcal{P}(\mathbb{R})$ of measures on $\mathbb{R}$ is endowed with the Monge-Kantorovich distance,

$$\mathbf{d}(m, m') = \sup_{\Xi} \int_{\mathbb{R}} \Xi(y) \mathrm{d}(m - m')(y), \tag{OA.9}$$

where the sup term is taken over all Lipschitz continuous maps $\Xi : \mathbb{R} \mapsto \mathbb{R}$ with a Lipschitz constant bound by 1.[59] Then the formal definitions are provided below, which also follow Cardaliaguet et al. (2019).

Definition 1. *We say that $J : \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ is C^1, if there is a continuous map $\frac{\delta J}{\delta m} : \mathcal{P}(\mathbb{R}) \times \mathbb{R} \to \mathbb{R}$ such that, for any $m, m' \in \mathcal{P}(\mathbb{R})$,*

$$\lim_{s \to 0^+} \frac{J((1-s)m + sm') - J(m)}{s} = \int_{\mathbb{R}} \frac{\delta J}{\delta m}(m, y) \mathrm{d}(m' - m)(y). \tag{OA.10}$$

Definition 2. *We say that $J : \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ is C^2, if for a fixed $y \in \mathbb{R}$, the map $m \mapsto \frac{\delta J}{\delta m}(m, y)$ is C^1. Moreover, denote its derivative as $\frac{\delta^2 J}{\delta m^2} : \mathcal{P}(\mathbb{R}) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. It satisfies*

$$\frac{\delta J}{\delta m}(m', y) - \frac{\delta J}{\delta m}(m, y) = \int_0^1 \int_{\mathbb{R}} \frac{\delta^2 J}{\delta m^2}\Big((1-s)m + sm', y, y'\Big) \mathrm{d}(m' - m) y' \mathrm{d}s. \tag{OA.11}$$

[59] As the online Appendix of Achdou et al. (2022) also discusses, in the theoretical literatures, the space is often specified as the n-dimensional torus $\mathbb{T}^n$ rather than $\mathbb{R}^n$. The only reason is to sidestep the discussion of boundary conditions in the space dimension.

Derivation. We adopt the second logic mentioned in Appendix OA3.1, which shares the similar idea with Bilal (2023). The key observation is to apply the Itô's Lemma to the value function J_t and especially note the time-dependence of m_t. As the time-dependence of w_t and A_t have been accounted in the HJB equation, the rest time-dependence could be replaced by the dependence on the distribution m.

$$\mathrm{d}J_t = \langle \frac{\delta J}{\delta m}, \mathrm{d}m_t \rangle + \frac{1}{2}\langle \mathrm{d}m_t | \frac{\delta^2 J}{\delta m^2} | \mathrm{d}m_t \rangle, \tag{OA.12}$$

where the inner product is defined in the appropriate functional space, $\langle f(x), g(x)\rangle = \int f(x)g(x)\mathrm{d}x$. $\langle h|f(x,x')|g\rangle = \langle\langle f(x,x'), h(x)\rangle, g(x')\rangle$. Substituting into the HJB equation, we obtain

$$\begin{aligned}\phi J(w,m,A)\mathrm{d}t = &\left\{\begin{array}{l} \mathcal{U}(w,m) + L(w)[J] + \mu^A A\dfrac{\partial J}{\partial A} + \dfrac{\partial}{\partial w}g(w)v(w) \\ \quad + \displaystyle\int \frac{\delta J}{\delta m}(w,w',m)L^*(w')[m]\mathrm{d}w' \\ \quad + \dfrac{1}{2}\displaystyle\int\int \frac{\delta^2 J}{\delta m^2}(w,w',w'',m)\frac{\partial g(w')m(w')}{\partial w'}\frac{\partial g(w'')m(w'')}{\partial w''}\mathrm{d}w'\mathrm{d}w'' \end{array}\right\}\mathrm{d}t \\ &- \left(v(w) + \int \frac{\delta J(w,w',m)}{\delta m}\frac{\partial g(w')m(w')}{\partial w'}\mathrm{d}w'\right)\mathrm{d}Z_t.\end{aligned} \tag{OA.13}$$

By the definition of the penalty, v guarantees that J is adapted with respect to the filtration generated by the shock, i.e., the same idea as Lemma 1, v lets the diffusion term zero. Substituting into (OA.13), the master equation (25) is derived.

Linking with the derivation process in literature. As mentioned in Section 4, a more standard and rigorous way to derive the master equation in MFG theory papers, e.g., as Cardaliaguet et al. (2019) does, is to make a change of variable, $\tilde{J}(w) = J(w + g(w)Z_t)$, $\tilde{m}(w) = m(w + g(w)Z_t)$,[60] implying that $\tilde{m}$ is the conditional law of the motion of the wealth process $w_t - g(w_t)Z_t$. When Z_t is relatively small, there is an approximation:

$$\mathrm{d}(w - g(w)Z_t) = [f(w) - Z_t g'(w)\mathrm{d}w]\mathrm{d}t \approx f(w)\mathrm{d}t.$$

Then $\tilde{m}_t$ satisfies:

$$\mathrm{d}\tilde{m}_t = -\frac{\partial}{\partial w}\left(\tilde{f}(w,\tilde{m})\tilde{m}_t(w)\right)\mathrm{d}t, \tag{OA.14}$$

where $\tilde{f}(w) = f(w + g(w)Z_t)$, and the tilde scripts below indicates similar meanings w.r.t. different functional without specific instructions. The HJB equation similarly becomes

$$\phi\tilde{J}_t\mathrm{d}t - \mathrm{d}\tilde{J}_t = \left[\tilde{\mathcal{U}}(w,\tilde{m}) + \tilde{f}(w,\tilde{m}) + \mu^A A\frac{\partial \tilde{J}_t}{\partial A}\right]\mathrm{d}t + \tilde{v}\mathrm{d}Z_t = \left[\tilde{\mathcal{U}}(w,\tilde{m}) + \tilde{f}(w,\tilde{m}) + \mu^A A\frac{\partial \tilde{J}_t}{\partial A}\right]\mathrm{d}t + \mathrm{d}\tilde{M}_t, \tag{OA.15}$$

where $\tilde{v}_t(w) = \tilde{g}(w)\partial_w\tilde{J} + v_t(w + g(w)Z_t)$. Denote the realization of shocks generated by $(Z_t)_{t\geq 0}$ as an information set $(\mathcal{F}_t)_{t\geq 0}$. Then $(\tilde{M}_t(w))_{t\in[0,T]}$ is an $(\mathcal{F}_{t\in[0,T]})$ martingale. (24), the boundary conditions $\tilde{m}_0 = m_0$, and $\tilde{J}_\infty$ bounded,

[60] As Cardaliaguet et al. (2019) illustrates, the formally equivalent but more rigorous definition of $\tilde{m}$ is to let it be a push-forward of m_t by the shift $w \mapsto w - g(w)Z_t$. This definition is completely licit. m_t reads as a conditional measure given the aggregate shock. Therefore, since conditioning consists of freezing the shock Z_t, such a shift is considered as a "deterministic" mapping.

jointly construct the standard expression of the MFG with aggregate shocks.

Discussion on the solution to the master equation. The classical solution to the second order master equation is, naturally, a map $J(w,m,A) : \mathbb{R}\times\mathcal{P}(\mathbb{R})\times\mathbb{R}\mapsto\mathbb{R}$, which satisfies all the requirements in the relevant definition part in this section, and satisfies the master equation. The definition of the weak solution introduced by Cardaliaguet and Souganidis (2021) focuses on the master equation with no idiosyncratic shock, i.e., the case in our paper.[61] The relationship between the weak solution and the classical solution is that if J^* is a weak solution, and J, $\frac{\partial J}{\partial w}$, $\frac{\delta J}{\delta m}$, $\frac{\partial^2 J}{\partial w^2}$, $\frac{\delta^2 J}{\delta m^2}$ and $\frac{\partial}{\partial w}\frac{\delta}{\delta m}J$ are continuous in w and m. Then J^* is a weak solution when the classical solution $J^*\ *$ can be expressed as J^* combining with a continuous function of m, $c^*(m)$, i.e.,

$$J^*(w,m,A) = J^{**}(w,m,A) + c^*(m). \tag{OA.16}$$

It is because the definition of the weak solution actually characteristics $\frac{\partial J}{\partial w}$ and not J. Under proper assumptions, the existence and uniqueness of the weak solution have been proved in literature.

OA3.3 Implications Under Heterogeneity

Consider the implications, i.e., the propositions in Section 3, under heterogeneity. Similar with the homogeneous case, the aggregate equilibrium value of tradable tokens is

$$\int_i x_i^* \mathrm{d}i = \left(\frac{1-\alpha}{r_t - c_t}\right)^{\frac{1}{\alpha}} N_t A_t \int_{\tilde{W}} \bar{U}_t(w_t) m(w_t) \mathrm{d}w_t \equiv \left(\frac{1-\alpha}{r_t - c_t}\right)^{\frac{1}{\alpha}} N_t A_t M_{1t}, \tag{OA.17}$$

where $\bar{U}_t = \mathbb{E}[U(u_{i,t}, w_t)|w_t]$ is a function of w_t, $\arg_w\{\bar{U}(w) > U_0\} = \tilde{W} \subset W$. Then $\Theta^* = 1 - \left(\frac{1-\alpha}{r-c}\right)^{\frac{1}{\alpha}} \frac{NAM_1}{q^*} = \Theta^*(r)$, where r fits the fixed point problem (11). Then we can see that the rest of proof are the same as Propostion 2 but only replace $\bar{U}$ with M_1. It is with no extra difficult to prove (17) under heterogeneity. Further, note that the additional terms in (25) does not enter the optimization, also the optimal q^* is not involved in extra terms. The derivation of pricing DE then still starts from the market clearing condition (further settings in Section 3.2 are adopted):

$$\begin{aligned} 0 &= \mu + r(\Theta) - c + \frac{\partial \Psi}{\partial n} + PQ_1\sigma^2, \\ 0 &= (1-\alpha)\int_{\hat{W}} \left(\frac{NA\bar{U}(w)}{q(w)}\right)^{\alpha} \frac{\partial_w J}{\partial_{ww} J} m(w)\mathrm{d}w + (\mu + \frac{\partial \Psi}{\partial n})\int_{\hat{W}} \frac{\partial_w J}{\partial_{ww} J} m(w)\mathrm{d}w + PQ_2\sigma^2, \end{aligned} \tag{OA.18}$$

where $Q_1 + Q_2 = Q$, $\hat{W} = W \setminus \tilde{W}$. The second equation refers to the aggregation of agents with zero staking. Denote the two integrals as $M_2 = M_2(A;m)$ and $M_3 = M_3(A;m)$, respectively, and sum up the two equations, we obtain

$$0 = (\mu + \Psi_n)(1 + M_3(A)) + r(\Theta) - c + (1-\alpha)M_2(A) + PQ/S\sigma^2. \tag{OA.19}$$

[61] The solution is solved by the Hilbert space approach introduced by Lions (2011), the detailed notion of the weak solution is provided in Cardaliaguet and Souganidis (2021). Here we no longer repeat existing work.

Then, apply Itô's Lemma to P_t and substitute the resulting (μ, σ) into the above equation, the pricing PDE is obtained and is similar with (20) but only involves M_s, $s = 1, 2, 3$ (M_3 enters Θ). The rest process is all similar with the baseline. Finally, $V(A)$ still satisfies an ODE. Further, M_2 and M_3 enters the "flow" term of the PDE, whereas M_1 is independent of A. Therefore, they do not qualitatively affect the relationship between the drift and the staking ratio.

To generate an intuitive understanding, we consider two specific cases. First, when all the agents have "extra" wealth to invest in staking, then $\hat{W} = \emptyset$ and $M_2 = M_3 = 0$, then (OA.19) is in the same form as (18). Second, if the transaction need is also linear in wealth, e.g., $U(u_{i,t}, w_{i,t}) = u_{i,t} w_{i,t}$, then it is not difficult to show that for agents with wealth $w_{i,t} \in \hat{W}$, $q(w_{i,t}) = x(w_{i,t}) = c_0 w_{i,t}$, where $c_0 > 0$ is independent of w, and $\partial_w \left(\frac{\partial_w J}{\partial_{ww} J} \right) = 0$. Denote $\int_{\hat{W}} \mathbb{E}(u_{i,t}^{\alpha} | w) m(w) \mathrm{d}w = \hat{M}_2$, $\int_{\hat{W}} m(w) \mathrm{d}w = \hat{M}_3$, then the summation of the two equations in (OA.18) yields

$$0 = (\mu + \Psi_n)(1 + \hat{M}_3) + r(\Theta) - c + c_1 (1 - \alpha)(NA)^{\alpha} \hat{M}_2 + PQ/S\sigma^2. \tag{OA.20}$$

Finally, the violation of UIP is directly obtained rearranging agents' optimal control and the aggregation across populations, and even without the assumption of log utilities. For agents with wealth $w \in \hat{W}$, they do not participate in staking and thus their optimal control problems do not involve r, i.e. their marginal excess return is the price appreciation plus the marginal transaction convenience. For agents with wealth $w \in \tilde{W}$, $\lambda_t = -\frac{\partial \Psi}{\partial n} + q_t^* \sigma^2 I$, which is completely in the same form as the baseline. As an aggregation, the system excess return of staking is

$$\lambda_t = -\Psi_n + \sigma^2 \int_{\tilde{W}} q^*(w) I(w) m(w) \mathrm{d}w, \tag{OA.21}$$

where $I(w) = \frac{\partial_{wwJ}}{\partial_{wJ}}(w) > 0$, $q^*(w) \geq 0$, $-\Psi_n$ is the marginal convenience of numéraire and is positive.

OA3.4 Discussion on the Properties of Stationary Distribution

To generate a better understanding of the stationary $m(w)$ (6), we consider the case when $\beta \to 1$, i.e., the transaction needs tends to be linear in wealth. $\frac{w^{\beta-1}}{\beta-1} \to \log(w)$ and thus

$$m(w) \to \bar{m}(w) = \bar{c}_0 \frac{1}{w^{(\phi - k_2)\mathcal{V}}} = \frac{[(\phi - k_2)\mathcal{V} - 1] \underline{w}^{[(\phi-k_2)\mathcal{V}-1]}}{w^{(\phi-k_2)\mathcal{V}}}, \tag{OA.22}$$

where the upper bar $\bar{\cdot}$ indicates the case at the limit. $\bar{m}(w)$ is a Pareto distribution with the tail parameter $(\phi - k_2)\mathcal{V} - 1$.

According to the properties of the tail parameter, when $(\phi - k_2)\mathcal{V} \leq 1$, there is no stationary situation. Considering the economic meaning of k_2 and $\mathcal{V}$, it corresponds to the situations that the benefits from on-chain activities or price appreciation are so large that the whole crowd continues to earn money. When $0 < (\phi - k_2)\mathcal{V} - 1 \leq 1$, the stationary distribution has infinite expectation, implying that the wealthier earn more from beneficial on-chain activities and price appreciation. When $(\phi - k_2)\mathcal{V} - 1 > 1$, the distribution has a finite expectation. The general case can be seen as a slight transformation of the Pareto distribution, which could roughly have similar different situations that corresponding to different on-chain convenience and price fluctuations.

On the other hand, the Pareto case can be easily solved by substituting $\beta \to 1$ into (A.9). Then it indicates

that in expectation among all the types, the wealth corresponding to transaction benefits changes linearly. Then by letting $h = m/w^2$, we obtain an ODE, $0 = a_2 w^2 h''(w) - a_1 w h'(w) + a_1 h(w)$, which finally results in a Pareto p.d.f. Since it is widely known that when the wealth of all people changes linearly, the resulting distribution by the law of motion follows the Pareto distribution, the above property also shows the consistency of the form of the general case. Also, it is worth pointing out that the general distribution has no contradiction with existing studies. Here, each agent's wealth still changes linearly. However, the slope varies among the crowd, which is affected by the user type $u_{i,t}$, whereas the randomness of $u_{i,t}$ is not independent on wealth. This cause the aggregation among populations to perform a slight difference from Pareto distribution.

OA3.5 Extended Discussion on the Feedback Effect

In the beginning of our setup, we introduce the feedback effect, which captures the "ideal" efforts of staking in incentivizing participation of the consensus or increasing the security level, etc., and thus enters the growth dynamic of platform productivity. The main relevant interest is how does the feedback effect influence token pricing. In Section 3, we have obtained the general positive relationship between staking ratio and expected price drift, which are both functions of the platform productivity (and some additional static arguments). Interestingly, the positive relationship exists no matter whether there is a feedback effect, but leads to different magnitudes of impact. Such a comparison is meaningful, since the feedback effect is not necessarily present in every stakable project. Also, it is somewhat difficult to verify its existence and size in practice. Then, the comparison shows general implications of that holds despite the existence or absence of feedback effects.

Further, the different magnitudes of impact are explained by two forces, in which the feedback force is that higher staking ratio leads to a higher drift of platform productivity, and then enters the price drift according to Itô's Lemma. Knowing such an extra benefit, agents will be slightly tilted towards staking in the on-chain trade-off. Thus, we see from Figure 1 that the equilibrium staking ratio is different under the same platform productivity. Figure OA.2 (A) shows such a variation more clearly. On the other hand as Subplot (B) shows, the aggregate on-chain wealth naturally increases with platform productivity, which is a corollary of the positive drift. The higher staking ratio is associated with higher drift, implying that there is more wealth allocated in as the platform grows.

The above two findings are also meaningful for discussing the feedback effect. However, it may rely on more assumptions or relatively narrow parameter choices. For the staking ratio, the setup of agents should not be too simplified to allow them to be able to anticipate the feedback effect, i.e. the growing platform, on themselves. If the utility is logarithmic and the transaction need is linear, then there would be no incentive to hedge against the changing investment opportunities presented by the dynamics of A, also they would be too myopia to see the benefits of future A growth. This setup would result in two overlapping curves for the numerical simulation of Subplot (A). As for the impact on the aggregate allocation, it generally holds that V increases with A, and when A is sufficiently large, the allocated wealth with feedback effects is larger, shown in Subplot (B). In periods where A is small, i.e., generally the initial phase of the platform, however, we observe in numerical simulations that platforms with feedback effects may instead have relatively small total wealth. This echoes the similar logic with the first force, i.e., there is a seesaw effect between already allocated wealth and potentially entering wealth. This

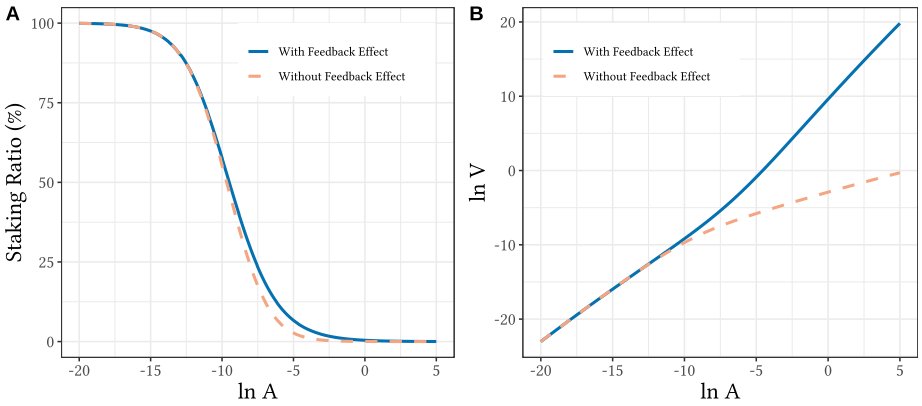


Figure OA.2: Platform productivity, staking ratio and aggregate on-chain wealth.
This plot shows how do the staking ratio and the aggregate on-chain wealth endogenously vary under different levels of platform productivity. The x-axis is the log productivity, $\ln A$. Subplot (A) and (B) show the corresponding curves of staking ratio and log on-chain wealth $\ln V$, respectively. The blue lines show the situations with feedback effects, whereas the orange dashed line are the situations without feedback.

is also comprehensively discussed in Cong et al. (2021b). The more in-depth discussion is however beyond the main focus of the present paper.

Overall, when staking is truly involved in building the platform rather than a purely speculative choice, with agents realizing that the platform development is beneficial to them, this feedback effect appears to work, and brings a greater influx of on-chain wealth to the platform, especially for well-developed platforms.

In addition, there are some different way to model the feedback effect. For example, since the growth of platforms may be accompanied by a broadening of usage scenarios, this effect can enter the differences between on-chain and off-chain convenience. It is worth pointing out that although we do some comparison the presence and absence of the feedback, however, our primary focus is still on how the on-chain actions are shaped in different platform states when there is such an effect.

OA3.6 Robustness in Sub-samples: Bulls and Bears

Table OA.1 A common concern related to pricing factors is the different predictability in bulls and bears. Fortunately, our data set includes a complete bull and bear market cycle. Overall speaking, the cryptocurrency market was roughly in a bull market in 2020 and 2021, and entered into a bear market in the end of 2021. A more precise and simple way to classify market bulls and bears refers to Lunde and Timmermann (2004) with the corresponding amplitude thresholds set to 35% (bulls) and 25% (bears). We use Bitcoin price as the indicator of the market, which is a common approach in practical investments. Also, as Ethereum is playing an increasingly important role, especially with its significant place in the staking economy, we alternatively use Ethereum price as the indicator for robustness. Figure OA.3 visualizes the segmentation of bulls and bears during the whole in-sample periods. We then regress the main empirical specification in Table 4 on sub-samples of bulls and bears. Columns (1)-(8) of Table OA.1 reports the regression results. In general, the estimated coefficients of $StakingRatio_{i,t-1}$ are all positive across bulls and bears that divided by different indicators, as well as different horizons (daily and weekly). This suggests that the implication that staking ratio positively predicts price appreciation generally holds in bulls and bears. It is worth noting that, however, the estimated value and significance

are lower in bears.

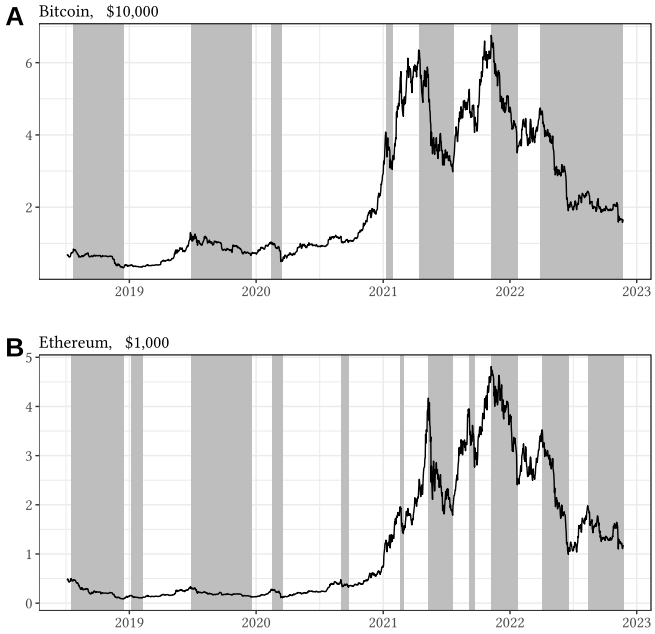


Figure OA.3: Identification of bull and bear markets.
Grey background denotes detected bear periods. In Subplot (A) and (B), we use the price of Bitcoin and Ethereum as indices, respectively. The identification algorithm refers to Lunde and Timmermann (2004) with the corresponding amplitude thresholds set to 35% (bulls) and 25% (bears).

Discussions on the proxy of platform productivity. The platform productivity is captured by the average onchain transaction processing per second, denoted as $a_{i,t}$ in empirical tests. In practice, numerous platforms and blockchains aim to increase the transaction size of the flows processed on their chains, reflecting the platform productivity. Meanwhile, it is also influenced by the overall transaction needs, which means it does not necessarily to measure the hardware upper limits. However, both the two forces are closely related to the concept of transaction convenience and fit our main use for inducing platform productivity. Therefore, it is not necessary here to separate the two forces. Note that it may not capture the whole progress of platform development, such as transaction security and performance on specific financial services. It is beyond our scope to suggest an aggregated measurement.

OA3.7 Robustness in Subsamples: PoS Tokens and DeFi Tokens

As mentioned in the introduction, our model evolve both base layer pan-PoS staking mechanisms and higher layer DeFi stakable tokens. We model the common features and introduce several implications. In the empirical analysis, we use a sample containing the tokens from both the two layers. To empirically illustrate that these implications are common for both the pan-PoS and DeFi tokens, we divide our sample into two subsets based on the category of tokens, and repeat the main tests of Table 4.

Columns (9)-(12) of Table OA.1 reports the results of these robustness tests. We regress pan-PoS and DeFi sub-samples on both daily and the weekly data sets. The dependent is log price change $r_{price_{i,t}}$ on the staking ratio in the previous week. The estimated coefficients of the staking ratio are both positive and consist with our

main empirical result, which suggest that the staking ratio predicts price appreciation. The daily data shows both significant estimations, whereas in the weekly data, the statistical power is lower. This also consists with the main regression in Table 4. Also, it could be partly explained by the small sample size since the raw weekly sample set is further divided into two subsamples.

Table OA.1: Robustness test for Table 4: Staking ratio and token prices.

This table presents the robustness test on the analysis of how the staking ratio predicts token price appreciation. The regression model is the same as the one used in Column (2) of Table 4, in which the main independent is the staking ratio of the previous period, $StakingRatio_{i,t-1}$, and the dependent $r_{price_{i,t}}$ is the log price change. The difference is that we replicate the test on subsets of bulls, bears, PoS tokens, and DeFi tokens. The bulls and bears are detected based on Lunde and Timmermann (2004)'s algorithm. Due to the lack of a recognized index for the cryptocurrency market, we refer to the common approach used in practice, i.e. using the Bitcoin price as a market indicator. We also repeat with Ethereum as an indicator for robustness. The detected bulls and bears are visualized in Figure OA.3. The subsets of pan-PoS and DeFi tokens are sorted based on tokens' nature. We also do the test in different horizons and with fixed effects to show the robustness of the results. Standard errors clustered in both token-specific and time dimensions are reported in parentheses. ***,**,* indicate statistical significance at the 1%, 5% and 10% respectively.

Dependent:	$r_{price_{i,t}}$											
	Bull-Bear								PoS-DeFi			
	Daily				7-Day				Daily		7-Day	
	Bitcoin as Indicator		Ethereum as Indicator		Bitcoin as Indicator		Ethereum as Indicator					
Sub-sample:	Bull	Bear	Bull	Bear	Bull	Bear	Bull	Bear	pan-PoS	DeFi	pan-PoS	DeFi
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$StakingRatio_{i,t-1}$	0.040***	0.010	0.025**	0.018	0.290***	0.017	0.135	0.124	0.028**	0.019*	0.143	0.167
	(0.009)	(0.012)	(0.011)	(0.012)	(0.098)	(0.094)	(0.091)	(0.109)	(0.012)	(0.010)	(0.107)	(0.097)
$\hat{\beta}_{i,t}$	-0.002	0.000	0.004	-0.011**	-0.051	-0.016	0.034	-0.169*	-0.003	-0.002	-0.016	-0.060
	(0.003)	(0.004)	(0.003)	(0.005)	(0.043)	(0.025)	(0.021)	(0.082)	(0.003)	(0.003)	(0.031)	(0.039)
$\log(Cap)_{i,t-1}$	-0.006***	-0.006***	-0.004***	-0.009***	-0.044***	-0.050***	-0.032***	-0.069***	-0.004***	-0.010***	-0.033***	-0.044*
	(0.002)	(0.002)	(0.001)	(0.003)	(0.012)	(0.016)	(0.008)	(0.020)	(0.001)	(0.002)	(0.009)	(0.023)
$r_{price_{i,t-1}}$	0.085*	-0.055	0.019	0.044	-0.116**	-0.004	-0.053	-0.195	0.066	0.098**	0.002	-0.112**
	(0.049)	(0.098)	(0.027)	(0.130)	(0.045)	(0.075)	(0.039)	(0.125)	(0.046)	(0.040)	(0.044)	(0.052)
$\Delta Network_{i,t-1}$	0.203***	0.353*	0.198***	0.412**	0.524	-0.284	0.421	-0.107	0.158***	0.259**	0.324	0.649**
	(0.062)	(0.178)	(0.065)	(0.190)	(0.314)	(0.332)	(0.262)	(0.357)	(0.047)	(0.111)	(0.343)	(0.290)
$a_{i,t-1}$	0.053	0.151	0.052*	0.114	0.448*	0.224	0.230	0.253	0.049	0.940***	0.276	1.145
	(0.035)	(0.116)	(0.027)	(0.121)	(0.245)	(0.771)	(0.198)	(0.895)	(0.044)	(0.227)	(0.317)	(2.280)
$Whale_{i,t-1}$	-0.018	0.006	-0.011	-0.011	-0.130	-0.107	-0.094	-0.141	-0.008	0.004	-0.073	-0.022
	(0.012)	(0.022)	(0.013)	(0.016)	(0.120)	(0.139)	(0.096)	(0.125)	(0.006)	(0.024)	(0.042)	(0.193)
$NotLaunched_{i,t}$	0.031***	0.001	0.021***	0.016	0.286***	0.075***	0.151***	0.270		0.006		0.142**
	(0.007)	(0.005)	(0.005)	(0.014)	(0.071)	(0.025)	(0.043)	(0.168)		(0.007)		(0.055)
$Y^0_{i,t}$	0.007*	0.007	0.005	0.012	0.078**	0.046	0.038	0.115	0.010*	0.007*	0.099**	0.079**
	(0.004)	(0.006)	(0.003)	(0.008)	(0.027)	(0.049)	(0.025)	(0.070)	(0.004)	(0.004)	(0.044)	(0.031)
Token FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,692	4,299	6,204	3,787	795	639	897	537	5,529	6,173	793	885
R^2	0.382	0.577	0.406	0.529	0.477	0.546	0.535	0.496	0.579	0.547	0.576	0.548

OA3.8 Staking Ratio and Returns

The predictability of price appreciation suggests that high staking ratio tokens should bring excess returns over the low ones. We test it by creating a long-short strategy, that is, sort the tokens by their staking ratio by the end of previous period, equal-weighted long the top 50% and short the bottom. The allocation is adjusted every week. Figure OA.4 and Table OA.2 document that the portfolio provides relatively stable positive cumulative returns with a Sharpe ratio of 0.865. To show that this is not another manifestation of the size effect, we test the same sorted strategy within the large-cap and small-cap token groups, respectively. The implication remains qualitatively robust. Table OA.2 suggests that the large-cap group performs weaker. It is partly because the staking ratio of these tokens are relatively low in general, and thus with smaller differences among each other. The corresponding effect is therefore minor.[62] Considering the limitations on shorting in practice, we also test the long-only strategies, i.e., borrow US dollars and equal-weighted long top (bottom) 50% tokens sorted by staking ratio. The top group outperforms the bottom group, and also the full-sample benchmark.[63]

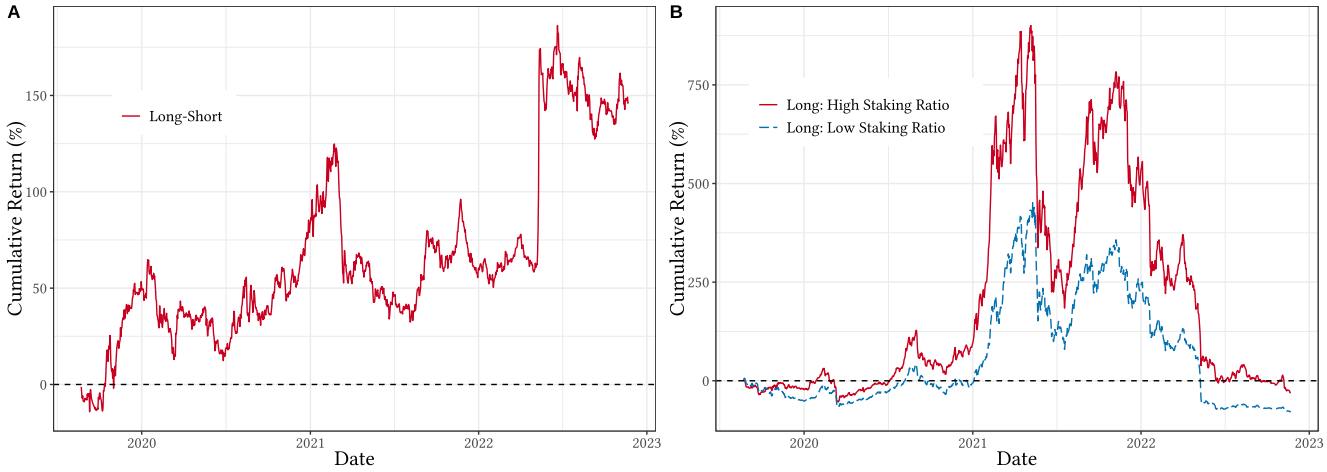


Figure OA.4: Cumulative returns of staking ratio sorted portfolios.
Panel A shows the long-short portfolio cumulative return. Panel B shows the long-only strategies. The portfolios are constructed as Section 6.2 describes.

OA3.9 Additional discussion on UIP and Crypto Carry

UIP with in-sample tokens as local currencies. Table OA.3 repeats the specification of Table 5, implying that the UIP is generally violated within the token market, rather than only comparing with mainstream cryptocurrencies or fiat moneys.

Summary statics of crypto carry. Table OA.4 summarizes the annualized carry and excess return of all tokens in our sample. Sample means and standard deviations are reported. We also include the US Dollar as one of the assets for which the carry and excess return are, by definition, equal to zero.

Crypto carry trade: cumulative returns. We visualize the cumulative returns of mentioned strategies in Figure OA3.9. Panel (a) shows the long-short strategies. The red curve in Figure plots the cumulative return of such a

[62] We also use Bitcoin as base asset

[63] We repeat this test using bitcoin denomination to cancel out the crypto market trend, the main results in Panel A remains robust.

Table OA.2: Sorted Portfolio performance by staking ratio.

This table reports the statistics of portfolio performance. The upper panel reports the results of the long-short carry strategy, including long top 50% high staking tokens and short bottom 50% across the full sample, i.e., corresponding to Figure OA.4, and same strategies but within top 50% large-cap and small-cap groups, respectively. The lower panel reports equal-weighted long strategies, including the full-sample benchmark, long top 50% high staking ratio tokens, and long bottom 50%. The portfolios are rebalanced each week. For each strategy, the annualized mean, standard deviations, skewness, kurtosis, maximum drawdown (MDD) and Sharpe ratio are reported.

Strategy	Mean (Annual, %)	St.dev. (Annual, %)	Skewness	Kurtosis	MDD (%)	Sharpe Ratio (Annual)
Long-short Strategy:						
Full Sample	36.081	41.731	2.273	27.582	41.135	0.865
Within Large-Cap Group	18.575	56.857	5.881	104.942	70.057	0.327
Within Small-Cap Group	40.052	62.720	0.588	5.177	40.065	0.639
Long Strategy:						
EW All assets	15.577	78.244	−1.576	7.672	92.934	0.199
EW High-Staking Ratio	22.873	80.737	−1.161	5.042	93.013	0.283
EW Low-Staking Ratio	−13.207	79.823	−1.724	9.547	96.115	−0.165

Table OA.3: UIP violation: in-sample local currencies.

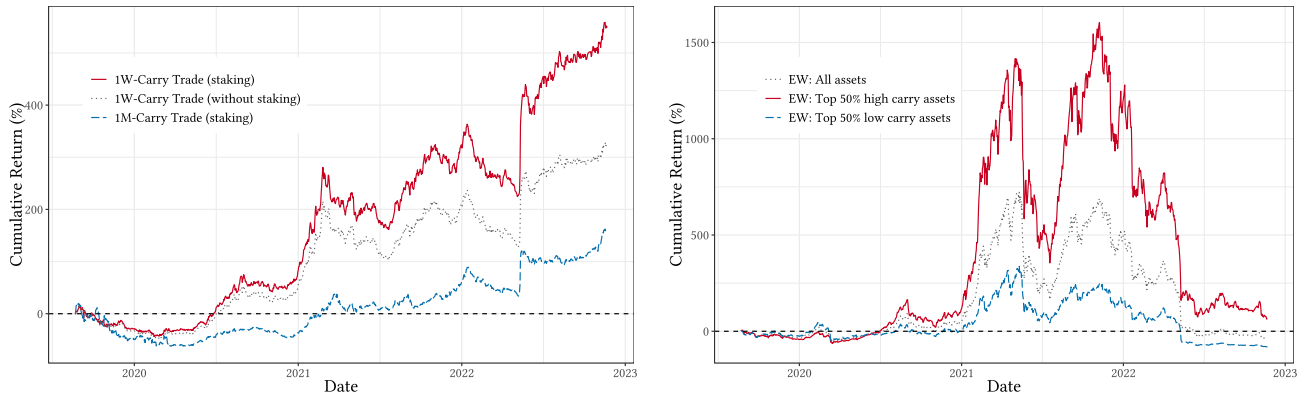
This table tests **(H3a)**, i.e., the panel regression results of UIP test. In each row, we use a different asset as local currency and report the estimated coefficients and corresponding robust standard errors clustered by tokens of β.

Local	7-day			30-day			Local	7-day			30-day		
Currency	Coef., β	Std. Err.	R^2	Coef., β	Std. Err.	R^2	Currency	Coef., β	Std. Err.	R^2	Coef., β	Std. Err.	R^2
Currency & mainstream cryptocurrencies.													
US Dollar	−1.02	(0.044)	0.33	−1.12	(0.176)	0.11	Ethereum	−1.04	(0.033)	0.37	−1.09	(0.126)	0.12
Bitcoin	−1.02	(0.034)	0.37	−1.08	(0.134)	0.11							
In-sample Tokens.													
1inch	−0.92	(0.039)	0.54	−0.72	(0.159)	0.22	kyber	−1.03	(0.024)	0.25	−1.01	(0.086)	0.14
aave	−1.03	(0.029)	0.32	−1.14	(0.119)	0.13	livepeer	−0.92	(0.013)	0.51	−0.52	(0.053)	0.24
aion	−0.90	(0.102)	0.43	−0.24	(0.604)	0.22	lto	−0.95	(0.129)	0.44	−0.16	(0.331)	0.25
algorand	−1.00	(0.014)	0.50	−1.12	(0.132)	0.13	matic	−1.06	(0.021)	0.37	−1.12	(0.086)	0.14
ark	−1.00	(0.053)	0.52	−0.90	(0.196)	0.30	mina	−1.06	(0.026)	0.30	−0.39	(0.224)	0.10
avalanche	−1.13	(0.049)	0.31	−1.34	(0.195)	0.13	mirror	−0.99	(0.008)	0.67	−0.88	(0.034)	0.32
band	−1.01	(0.031)	0.32	−1.05	(0.128)	0.11	near	−1.10	(0.033)	0.32	−1.42	(0.161)	0.14
bifi	−1.01	(0.067)	0.34	−0.80	(0.308)	0.16	nem	−1.07	(0.039)	0.58	−1.29	(0.330)	0.34
binance-sc	−0.99	(0.019)	0.36	−1.00	(0.081)	0.12	neo	−0.86	(0.053)	0.53	−0.47	(0.289)	0.26
bitbay	−0.83	(0.068)	0.17	−0.40	(0.790)	0.07	nuls	−0.98	(0.077)	0.49	−1.30	(0.602)	0.17
cardano	−1.01	(0.027)	0.33	−1.02	(0.134)	0.10	oasis	−1.04	(0.034)	0.29	−1.12	(0.135)	0.12
cosmos	−1.06	(0.035)	0.37	−1.17	(0.139)	0.13	olympus	−1.00	(0.010)	0.83	−1.12	(0.034)	0.49
cronos	−1.03	(0.040)	0.29	−1.07	(0.172)	0.22	osmosis	−0.91	(0.028)	0.41	−0.48	(0.106)	0.25
curve	−1.05	(0.028)	0.33	−1.15	(0.120)	0.13	pancakeswap	−1.13	(0.018)	0.74	−1.44	(0.070)	0.51
dash	−0.88	(0.096)	0.42	−0.56	(0.343)	0.24	peakdefi	−1.02	(0.027)	0.54	−1.03	(0.102)	0.25
decred	−1.03	(0.034)	0.37	−1.05	(0.123)	0.11	polkadot	−1.05	(0.031)	0.35	−1.69	(0.275)	0.08
dfinity	−0.98	(0.019)	0.29	−0.84	(0.120)	0.17	qtum	−0.87	(0.036)	0.57	−0.56	(0.156)	0.30
dodo	−0.91	(0.027)	0.33	−0.57	(0.149)	0.14	secret	−1.13	(0.030)	0.32	−1.68	(0.180)	0.14
dydx	−0.91	(0.050)	0.27	−0.77	(0.122)	0.29	smartcash	−0.90	(0.070)	0.45	−0.51	(0.403)	0.25
elrond	−1.06	(0.028)	0.33	−1.19	(0.126)	0.14	snx	−1.05	(0.018)	0.61	−1.20	(0.086)	0.31
eos	−1.10	(0.033)	0.50	−1.36	(0.142)	0.23	solana	−1.09	(0.052)	0.32	−1.17	(0.182)	0.12
eth2.0	−1.04	(0.014)	0.42	−1.25	(0.059)	0.17	stafi	−1.01	(0.030)	0.31	−1.10	(0.158)	0.12
fantom	−0.89	(0.013)	0.47	−0.48	(0.065)	0.10	stake-dao	−0.40	(0.087)	0.30	2.15	(0.323)	0.25
flow	−1.04	(0.027)	0.27	−1.03	(0.102)	0.12	sushi	−1.15	(0.085)	0.44	−1.16	(0.117)	0.15
harmony	−1.07	(0.032)	0.30	−1.31	(0.134)	0.12	terra	−2.10	(0.423)	0.05	−4.52	(1.686)	0.05
icon	−0.91	(0.071)	0.47	−0.70	(0.491)	0.22	tezos	−1.00	(0.028)	0.36	−0.93	(0.122)	0.11
idex	−1.11	(0.047)	0.28	−1.21	(0.121)	0.12	tron	−1.02	(0.033)	0.36	−1.05	(0.134)	0.11
injective	−1.06	(0.034)	0.27	−1.14	(0.157)	0.12	wanchain	−0.93	(0.073)	0.49	−0.61	(0.415)	0.27
iotex	−0.91	(0.029)	0.46	−0.63	(0.229)	0.22	waves	−1.01	(0.033)	0.31	−0.80	(0.148)	0.09
irisnet	−1.04	(0.046)	0.43	−0.97	(0.179)	0.18	wax	−1.02	(0.028)	0.26	−1.09	(0.114)	0.09
kava	−1.11	(0.018)	0.50	−1.41	(0.072)	0.24	yearn	−1.06	(0.021)	0.52	−1.14	(0.075)	0.18
kusama	−1.03	(0.031)	0.31	−1.10	(0.144)	0.12	zcoin	−0.87	(0.055)	0.57	−0.77	(0.316)	0.31

Table OA.4: Excess return and carry.

Token	Excess Return (%, Annual)		Carry (%, Annual)		Token	Excess Return (%, Annual)		Carry (%, Annual)	
	Mean	Std.dev.	Mean	Std.dev.		Mean	Std.dev.	Mean	Std.dev.
0x	0.42	0.30	6.75	20.62	kusama	12.97	1.93	16.61	25.75
1inch	2.94	6.49	1.46	17.05	kyber	1.32	3.53	1.88	17.67
aave	3.83	0.98	6.21	21.56	livepeer	62.04	29.54	63.61	49.84
aion	6.25	3.01	8.86	18.61	lto	6.70	1.01	12.19	17.18
algorand	7.20	11.72	7.81	19.22	matic	17.78	14.01	25.23	40.64
ark	8.09	0.52	9.52	17.02	mina	10.32	2.81	9.60	20.97
avalanche	8.43	2.74	13.31	35.40	mirror	39.15	37.41	34.85	42.27
band	10.88	3.10	13.62	27.93	near	10.13	2.85	13.30	23.35
bifi	7.95	3.43	11.76	26.02	nem	−1.35	0.51	−1.59	14.28
binance-sc	8.02	6.38	9.05	15.10	neo	0.92	0.97	2.71	14.94
bitbay	1.13	0.98	10.62	63.06	nuls	8.31	0.56	10.67	16.96
cardano	4.39	2.77	6.52	17.95	oasis	11.80	4.55	12.95	26.77
celo	6.08	0.13	6.07	4.95	olympus	49.79	41.90	37.95	48.49
cosmos	9.82	2.35	11.97	18.55	osmosis	35.90	17.56	30.03	20.48
cronos	10.15	2.67	6.25	13.11	pancakeswap	74.76	26.66	78.13	51.65
curve	1.12	2.83	1.45	19.34	peakdefi	43.88	16.92	43.26	32.91
dash	5.20	0.73	7.28	26.50	polkadot	11.56	1.68	13.14	17.63
decred	5.58	1.71	6.66	14.32	qtum	4.73	1.31	6.30	14.07
dfinity	7.68	4.78	3.87	15.34	secret	24.47	3.95	27.70	28.47
dodo	56.63	10.73	50.54	22.66	smartcash	1.63	0.36	3.64	15.35
dydx	10.66	2.76	8.53	19.03	snx	21.96	23.54	26.42	37.33
elrond	14.24	7.45	19.18	32.93	solana	5.94	3.82	8.94	24.75
eos	10.69	12.06	11.02	18.81	stafi	18.76	4.02	19.91	27.21
eth2.0	8.61	10.82	11.45	18.44	stake-dao	22.23	8.32	21.49	19.43
fantom	27.83	23.95	37.98	47.03	sushi	10.51	10.12	8.85	20.06
flow	6.95	2.05	3.29	16.56	terra	8.26	3.71	14.26	35.66
harmony	8.58	2.89	12.42	27.40	tezos	4.56	2.11	5.52	16.99
icon	16.42	2.80	19.99	23.99	tron	2.81	1.94	4.36	13.99
idex	8.05	8.85	15.12	78.53	wanchain	7.39	0.26	9.44	16.86
injective	3.87	0.58	2.16	13.91	waves	3.84	1.62	6.21	22.75
iotex	8.86	3.14	11.27	21.23	wax	1.56	2.64	3.34	19.81
irisnet	9.67	0.38	14.77	22.07	yearn	14.51	16.78	16.86	30.14
kava	19.55	16.44	22.29	26.61	zcoin	15.02	3.77	18.35	16.90
US Dollar	0.00	0.00	0.00	0.00					

carry trade strategy. It shows an overall increase and large cumulative returns. Especially, in the cryptocurrency market where price volatility is huge, such a strategy performs a relatively smooth growth, implying the carry premia always exists. The gray line shows the performance of the same carry portfolio, but without staking. That is, for long tokens, we do not stake them, and for short assets, we also do not compensate for the staking reward rate. The strategy also exhibits increasing cumulative returns, which implies that the carry strategy earns excess returns not only from carry (staking reward) but also from price appreciation. Moreover, the blue line reports the performance of the same carry portfolio, but is rebalanced every month. It exhibits fewer returns than 1W-carry trade. There are two potential explanations. To compare with the benchmark of the equal weighted long strategy, also consider the potential short-selling restrictions, we test the performance of the long-only strategies as Panel (b) shows. Since the market fluctuations are not hedged, all the strategies are volatile and move in co-trends. However, the strategy that go long top 50% tokens with high carry still provide a relatively better performance with a larger Sharpe ratio as Table 6 reports in the main text.



(a) Long-short strategies.

(b) Long strategies.

Figure OA.5: Cumulative returns of carry strategies.
This figure corresponds to Hypothesis **(H3b)** and shows the cumulative returns of the carry-trade strategies. Panel (a) shows the long-short strategies. The red curve shows the cumulative return of main long-short carry strategies. The construction of the strategy is described in Section 6.3 and $X = 50$. The choice of X does not affect the main qualitative properties. As comparisons, the gray curve corresponds to the strategy without earning or compensating staking rewards, and the blue curve shows the strategy that rebalanced every month. Panel (b) shows the long-only strategies. The gray curve corresponds to the equal-weighted benchmark, i.e., borrow US dollar and long all the tokens with equal weight. The red curve shows the result of the top 50% EW strategy, that is, borrow US dollar and go long top 50% high carry tokens with equal weight. The blue curve shows the result of the lowest 50% EW strategy, i.e., borrow US dollar and go long the 50% tokens with the lowest carry with equal weight. The order of the tokens is evaluated every week.

References for Online Appendices

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2022): "Income and wealth distribution in macroeconomics: A continuous-time approach," The review of economic studies, 89 (1), 45–86.

Augustin, Patrick, Roy Chen-Zhang, and Donghwa Shin (2022): "Yield Farming," Available at SSRN 4063228.

Bertucci, Charles (2021): "Monotone solutions for mean field games master equations: continuous state space and common noise," arXiv preprint arXiv:2107.09531.

Bilal, Adrien (2023): "Solving heterogeneous agent models with the master equation," Tech. rep., National Bureau of Economic Research.

Brunnermeier, Markus K and Yuliy Sannikov (2016): "On the optimal inflation rate," American Economic Review, 106 (5), 484–89.

Cardaliaguet, Pierre (2010): "Notes on mean field games," Tech. rep., Technical report.

Cardaliaguet, Pierre, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions (2019): "The master equation and the convergence problem in mean field games," in The Master Equation and the Convergence Problem in Mean Field Games, Princeton University Press.

Cardaliaguet, Pierre and Charles-Albert Lehalle (2018): "Mean field game of controls and an application to trade crowding," Mathematics and Financial Economics, 12 (3), 335–363.

Cardaliaguet, Pierre and Panagiotis Souganidis (2020): "On first order mean field game systems with a common noise," arXiv preprint arXiv:2009.12134.

——— (2021): "Weak solutions of the master equation for Mean Field Games with no idiosyncratic noise," arXiv preprint arXiv:2109.14911.

Carmona, Rene (2020): "Applications of mean field games in financial engineering and economic theory," arXiv preprint arXiv:2012.05237.

Cong, Lin William and Zhiguo He (2019): "Blockchain disruption and smart contracts," Review of Financial Studies, 32 (5), 1754–1797.

Cong, Lin William, Zhiguo He, and Jiasun Li (2021a): "Decentralized mining in centralized pools," Review of Financial Studies, 34 (3), 1191–1235.

Cong, Lin William, Ye Li, and Neng Wang (2021b): "Tokenomics: Dynamic adoption and valuation," Review of Financial Studies, 34 (3), 1105–1155.

Harvey, Campbell R, Ashwin Ramachandran, and Joey Santoro (2021): DeFi and the Future of Finance, John Wiley & Sons.

Irresberger, Felix, Kose John, Peter Mueller, and Fahad Saleh (2021): "The public blockchain ecosystem: An empirical analysis," NYU Stern School of Business.

John, Kose, Thomas J Rivera, and Fahad Saleh (2020): "Economic implications of scaling blockchains: Why the consensus protocol matters," Available at SSRN 3750467.

Krusell, Per and Anthony A Smith, Jr (1998): "Income and wealth heterogeneity in the macroeconomy," Journal of political Economy, 106 (5), 867–896.

Lasry, Jean-Michel and Pierre-Louis Lions (2007): "Mean field games," Japanese journal of mathematics, 2 (1), 229–260.

Li, Zongxi, A Max Reppen, and Ronnie Sircar (2019): "A mean field games model for cryptocurrency mining," arXiv preprint arXiv:1912.01952.

Lions, Pierre-Louis (2011): "Courses at the Collège de France." .

Lunde, Asger and Allan Timmermann (2004): "Duration dependence in stock prices: An analysis of bull and bear markets," Journal of Business & Economic Statistics, 22 (3), 253–273.

Saleh, Fahad (2021): "Blockchain without waste: Proof-of-stake," Review of Financial Studies, 34 (3), 1156–1190.