Can Markets be Fully Automated?
Evidence From an 'Automated Market Maker'*

Peter O'Neill†

November 3, 2021

Link to Latest Version

Abstract

Would a *fully* automated financial market be desirable, or even possible? While algorithms are increasingly prevalent in modern markets, they are still managed by humans who intervene in the operation of the market. In contrast, a new market type, "Automated Market Makers" (AMMs), removes humans completely, automating their trade matching and liquidity provision functions in transparent and deterministic code. Can such markets succeed? Using a complete record of 39 million AMM transactions from blockchain data, I show that AMMs now provide liquidity to over $50bn worth of trades per month. Despite the crude simplicity of their pricing algorithm, AMMs facilitate price discovery and their prices manage to stay within close bounds of other less automated markets. The overcome adverse selection costs in liquidity provision through fees, making their liquidity sustainable. I show that these markets arrive at equilibrium levels of liquidity by adjusting the size of the liquidity pool, analogous to traditional market makers adjusting bid-ask spreads. Consequently, they provide deeper liquidity pools in less volatile assets. Lastly, I show that AMM liquidity is more stable during extreme volatility than traditional markets, suggesting that full automation might have benefits.

Keywords: Market Design, Liquidity, Automated Market Makers, Cryptocurrency, Uniswap

JEL Classification: D47, G14

*I appreciate comments from Luke Anderson, Eric Budish, Anne Dyhrberg, Martin D.D. Evans, Sean Foley and Talis Putnins.

†Financial Conduct Authority (UK). peter.oneill@fca.org.uk.
The views expressed are those of the author and do not necessarily necessarily represent the position of the FCA.

1 Introduction

Today, Fischer Black's 1971 vision of a fully automated exchange[1] is mostly a reality. 'Electronic market-makers' have replaced human 'dealers' with algorithms that decide the price and amount of liquidity supplied in markets.[2] But this automation is not complete. Humans manage algorithms, they intervene to improve their design, or turn them off entirely if conditions become unfavorable, leading to withdrawals of liquidity that may exacerbate market volatility.[3]

Suppose a market, instead, consisted of just a single algorithm for facilitating trades. One that was simple, static and deterministic. One that fulfilled both the trade matching function of markets and handled all of the liquidity provision. This algorithm couldn't be turned off. It would contain an ex-ante binding mechanism that creates certainty on both the quantity and cost of liquidity for traders. A commitment to provide liquidity without human intervention. This complete level of automation goes much further than Fischer Black ever envisioned, he may have considered even the notion entirely unworkable. Perhaps it is?

Would fully automating a financial market be desirable, or even possible? If the algorithm is unable to cancel or re-price resting orders perhaps its quotes are always stale, leading it to be 'sniped' by arbitrageurs and incur maximum adverse selection costs, (Budish et al., 2015). If it cannot revise prices to reduce the inventory it has on hand, it is exposed to maximum inventory risks. It would need a mechanism to ensure these risks are mitigated, or at least a mechanism to fully compensate for them. In compensating for these high risks, would the cost of trading be driven so high that no investor would want to trade in these markets?[4]

This paper examines the success of this experiment in full automation currently un-

[1] Black (1971a) and Black (1971b).

[2] On the impacts of automation: Hendershott et al. (2011) and Chaboud et al. (2014), and on automated market making firms: Menkveld (2013), Brogaard et al. (2015), Foucault et al. (2016), Shkilko and Sokolov (2020), Weller (2018) and Van Kervel and Menkveld (2019).

[3] See liquidity withdrawals in Anand and Venkataraman (2016), flash crashes in Kirilenko et al. (2017), extreme price movements in multiple securities in Brogaard et al. (2016), correlation of liquidity across securities in Malceniece et al. (2019), and increasing premiums on liquidity risk in Pastor and Stambaugh (2019).

[4] Glosten and Milgrom (1985) note that if information asymmetry is too severe uninformed trading will not occur and the market maker cannot obtain sufficient revenues to compensate for trading against informed traders, similar to Akerlof (1970).

derway in cryptocurrency markets. A simple formula, $x * y = k$, sets prices and quantities for thousands of different assets using only 378 lines of code.[5] Called 'Automated Market-Makers' (AMMs), they execute upwards of \$50bn USD per month in digital assets — around half of the value traded on the largest US cryptocurrency exchange, Coinbase. I examine 'Uniswap', the largest of the AMMs, using a year of data acquired directly from the public Ethereum blockchain containing over 39 million transactions.[6]

Anyone can become an Liquidity Provider (LP) to an AMM by transferring their digital assets to discrete 'pools'. They can add to the liquidity of existing pools, thereby increasing the pool size, or create new pools. In return they receive revenue from a fixed 30 basis point fee on trades (called 'swaps') by liquidity demanders to the AMM. Fee revenue is shared equally amongst LPs in proportion to their investment — so LP returns vary as a function of the amount of liquidity in a pool (or the pool's 'size'). LPs can only add or remove liquidity, prices are set by the AMM algorithm, unlike traditional market makers (MMs) that can also vary prices.

This paper demonstrates that full automation is possible. AMMs have an effective mechanism to overcome adverse selection costs by varying the *amount* of liquidity they provide (pool size) as the equilibrating mechanism in contrast to traditional MMs that vary the *cost* of liquidity (the spread). This equilibrating mechanism ensures sustainable liquidity provision in AMMs. I observe these equilibrium effects in a time-series analysis of pool-returns. Liquidity flows into more profitable pools and these flows then reduce fee yields. I also observe equilibrium effects in the cross-section: higher volatility pools are associated with higher fees, and larger pools experience less volatility and reduced order flow toxicity.

Using the AMMs pricing function, I derive and test the components that govern returns to LPs, reconciling them to established market microstructure concepts. I empirically show that LP returns can be expressed as a function of the underlying prices of the assets in the pool, as well as the balanced and unbalanced order flow. This orderflow imbalance is shown to be a significant source of variation in pool size and profitability.

I find that AMM prices closely match those in centralized markets after accounting for

[5] For the core contracts of Uniswap programmed in Solidity, see: github.com/Uniswap/v2-core

[6] First, the entire Ethereum blockchain is synchronized. Uniswap transactions and fields are then parsed from the raw blockchain using 'TheGraph' API.

arbitrage constraints. This implies that price discovery occurs in AMM markets. Efficient prices aid the feasibility of AMMs from a liquidity provision perspective, as if prices are consistently stale, adverse selection costs will be higher. Additionally, a large proportion of tokenized assets traded on Uniswap are not traded on centralized exchanges, so the majority (or all) of price discovery is occurring on AMMs in these tokens.

I provide evidence that fully automated markets improve both the stability and resiliency of liquidity. During extreme periods of volatility AMMs do not experience liquidity withdrawal in comparison to centralized exchanges and also experience significantly less volatility in prices during these periods. This improved resiliency may arise because prices are set on AMMs in periodic batches, the timing of which varies between 11 and 19 seconds — the duration between blocks on the Ethereum blockchain. This batching might prevent 'feedback loop price' dynamics on orderbook markets. Market-makers in AMMs are also charged a 'gas fee' to submit and remove liquidity, which is significantly larger than the fees to trade, perhaps encouraging less 'fleeting liquidity' behaviors.

This paper relates AMMs to classical theories of market making. These theories derive how MMs can optimally address three challenges of inventory control, adverse selection risks, and setting adequate compensation for these challenges. Further, they imply that suboptimal solutions to these problems will lead to an inefficient liquidity provider that is unlikely to survive competition from more efficient providers — an implication that potentially spells doom for poorly designed market making mechanisms. I contrast how AMMs solve those challenges through their pricing function and equilibrating mechanism.

The first challenge, inventory control, is ensuring that the MM's inventory quickly reverts back towards the desired level (e.g., a zero or "flat" position in the risky asset) after trades, because holding long or short inventory entails risk. Inventory control models show that MMs can solve this challenge by optimally "shading" quotes down after trading with a seller to attract buyers, and up after trading with a buyer to attract sellers, thereby reverting her inventory (e.g., Stoll (1978); Ho and Stoll (1981); Hendershott and Menkveld (2014)). AMMs, similarly have an implicit quote shading mechanism within their pricing formula — following a buy, the AMM's price offered to the next trader increases, incentivizing sellers to would help revert the AMMs inventory level, and vice versa. The difference is that rather than having an "optimal" quote shading function as derived by inventory control models, AMMs have an arbitrary function that moves prices in the direction implied by inventory control models but not necessarily by an optimal amount. Instead, the quote shading function in AMMs is designed so ensure the AMM cannot run out of inventory of either of the two assets in its pool. It does this by increasing the price of an asset at an increasing rate as the remaining inventory of that asset falls.

The second challenge is adverse selection risks. MMs learn about asset values from order flows to avoid making excessive losses to informed traders. Models of asymmetric information (e.g., Kyle (1985); Glosten and Milgrom (1985)) show that a MM facing a mix of informed and uninformed order flow can use Bayesian learning to optimally extract the private information from order flow and use it to update their beliefs about the asset value and adjust prices accordingly. This learning is what drives price discovery — prices converge towards their fundamental values, eventually revealing all private information. The models imply that learning the private information as quickly as possible is crucial for a MM to minimize their adverse selection losses to informed trades. In reduced form, Bayesian learning results in MMs increasing quotes following buys and decreasing them following sells, by an amount proportional to the informativeness of the order flow. I show that the simple deterministic pricing function of AMMs has similar reduced form behavior, with prices adjusting up/down following buys/sells. Even the amount of the price change following a trade is related to the informativeness of order flow because, as I show, pools tend to be smaller and thus price impacts larger for assets with more informed trading and higher volatility. Thus, AMMs function as if they have a mechanism to learn from order flow, albeit a very simple one, allowing price discovery to occur in AMMs. However, whether the rate of implied learning in an AMM makes their price discovery too inefficient for them to be competitive in liquidity provision is an empirical question explored in this paper.

While AMMs incorporate "private information" via trades, as in Kyle (1985); Glosten and Milgrom (1985), more recent models such as focus on adverse selection arising from "public information'. Budish et al. (2015) model a market maker that tries to revise their quotes in response to public information (information that is symmetrically observable) shocks before "snipers" are able trade on them.[7] AMMs differ from these models as they

[7] Others such as Foucault (1999), Foucault et al. (2003), Hoffmann (2014), Biais et al. (2015) also

are unable to revise quotes in response to public information, except via trading (i.e., having their stale quotes 'sniped').[8] The inability for AMMs to avoid public adverse selection may result in higher liquidity costs in AMMs in comparison to traditional MMs. Though, this may be offset by the lack of expenditure by LPs on the "arms race" for speed (as in Budish et al. (2015)) that public adverse selection avoidance entails.

Finally, the market making theories discussed above also derive the level of trading costs (e.g., bid-ask spread) charged by the MM to exactly cover the expected costs of market making, including an inventory holding risk premium and adverse selection costs. Competition drives bid-ask spreads towards these breakeven levels such that an inefficient MM that has not minimized these costs and requires a wider bid-ask spread will be driven out of the market by more competitive MMs. The models imply that equilibrium bid-ask spreads vary across assets because inventory holding risk and adverse selection costs are a function of asset properties such as volatility and the amount of informed/uninformed trading. I show that the same cross-asset variation in liquidity is true of AMMs, although the equilibrating mechanism is somewhat different. Rather than setting a bid-ask spread, AMMs have a price impact function and a fixed proportional fee per transaction. I show that the fee in AMMs serves the same role as the bid-ask spread in market marking theories — it allows LPs to recoup compensation for adverse selection costs and inventory holding risk. But, unlike bid-ask spreads, the fees in AMMs are fixed. Therefore, AMMs arrive at an equilibrium level of liquidity by varying the pool size so that a given amount of fee revenue is shared among fewer liquidity providers (higher fees per provider) when the costs of providing liquidity are higher (e.g., more volatile assets) and conversely for assets with lower liquidity provision costs (e.g., less volatile assets). As a result, AMM liquidity varies across assets much like what is implied by the market making theories, although with a

model the market maker's ability to revise stale quotes as an important driver of the cost of liquidity provision. Budish et al. (2020) extend their model to include both public and private information and Aquilina et al. (2021) show that this adverse selection derived from public signals is significant, at 33% of total adverse selection. It is worth noting, however, that this estimate is from a limit orderbook, which represents an equilibrium where the MM can revise quotes. In an AMM environment where they cannot, this will be higher. Brogaard et al. (2021) decompose stock returns driven by market-wide, public and private information. They find that the return variance is 37% of public firm-specific information, 24% private firm-specific 8% market-wide and the rest is noise.

[8] An AMM has recently launched called "DODO" which attempts to resolve public information arbitrage by introducing an external pricing feed to anchor the price between two assets in the AMM. This is demonstrably an incomplete solution however, as most traded pairs on "DODO" are entirely stable-coin tokens as the stability of their price relationships is highly certain, (DODO, 2020).

novel mechanism for determining the liquidity level.

Therefore, the simple pricing function $xy = k$, combined with a fixed fee and variable pool size, provides AMMs with the necessary mechanisms to address the three key theoretical challenges in market making — inventory control, price discovery by learning from order flow, and setting breakeven liquidity costs that vary with asset characteristics. What remains to be seen as an empirical question is whether the AMM mechanisms are also sufficient in these functions because the simple AMM design comes at the cost of not adhering to the theoretically optimal inventory control or learning mechanisms. The ultimate question is whether this "inefficiency" in the AMM's design will render its liquidity provision costs uncompetitive and therefore not sustainable in the face of competition, or whether the lure of simplicity outweighs the inefficiencies, which may be small in practice? My results suggest the latter is currently true, but whether this model survives as a viable new market design in the long-run is uncertain.

Research on AMMs is at an early stage and has focused mostly on limitations in the current design, rather than the extent to which they are successful or beneficial. Aoyagi (2020) develops a Glosten and Milgrom (1985) style model that demonstrates a stable equilibrium can occur between competitive LPs. Lehar and Parlour (2021) examine the price discovery and liquidity of Uniswap AMMs as compared to centralized exchanges. Park (2021) focuses on the role of arbitrageurs, showing that the design of constant product AMMs can be improved to reduce front-running on the blockchain network, as also documented in Daian et al. (2019). Capponi and Jia (2021) model an AMM alongside a centralized market, showing that liquidity provision in AMMs should increase in response to trading volumes, decrease in response to volatility and impose gas fee externalities on the blockchain network.[9]

[9] There are many other papers that examine AMMs from a computer science tradition. Xu et al. (2021) derives adverse selection functions and the conservation function for several competing AMMs, while the function for Uniswap V2 is simple ($x * y = k$) others are more complex. Angeris et al. (2019) formalizes mathematical properties of Uniswap, Angeris et al. (2021) proposes new AMM designs that replicate options payoff structures for LPs and Angeris et al. (2020) demonstrates that the curvature of the AMM conservation function can protect LPs from adverse selection in volatile assets.

2 How do AMMs Work?

AMMs allow users to buy and sell tokenized assets on a blockchain without the involvement of a centralized exchange such as 'Coinbase' or 'Binance'.[10] An AMM consists of code written as an Ethereum smart contract which facilitates trade according to a deterministic, static algorithm.

Uniswap is comprised of distinct liquidity 'pools' which each hold two assets (or tokens). The most popular pool has the tokenized form of USD (USDT) and tokenized Ethereum (ETH).[11] Consider x_t and y_t the quantities of USD and ETH in the AMM at time t, respectively.

Pools are created by individual liquidity providers (LPs) who transfer their balances in ETH and USDT to the pool's blockchain address. LPs provide (or 'stake') a total of x_0 of USD and y_0 of ETH at $t = 0$ when the price of ETH in dollars is P_0. This gives the pool the constant $K = x_0 y_0$. The total USD value of the assets in the AMM at $t = 0$ is V_0, which (as I show later) is equal to $2x_0$ or twice the amount of USD deposited in the AMM.

This creates resting liquidity which can then be accessed by liquidity demanders (LDs or "traders") by 'swapping' any one of the two assets for the other. A trade[12] is an asset swap in which one asset is added and the other removed in quantities such that the constant K is maintained (by design of constant product AMMs such as Uniswap). Thus, a trade at time t that buys ETH involves removing $\Delta y_t = y_t - y_{t-1}$ from the AMM and adding $\Delta x_t = x_t - x_{t-1}$ such that the constant k is preserved, $x_t y_t = K = x_{t-1} y_{t-1} = (x_{t-1} - \Delta x_t)(y_{t-1} + \Delta y_t)$. Conversely, a trade at time t that sells ETH by adding Δy_t to the AMM and removing Δx_t from the AMM, $x_t y_t = (x_{t-1} - \Delta x_t)(y_{t-1} + \Delta y_t)$.

Therefore, the price of the pool is a function of the ratio x and y, but also the size of

[10] This means AMMs are sometimes referred to as "decentralized exchanges" (DEXs). This is appealing to users as it fulfills the promise of blockchain technology not being reliant on traditional finance institutions and human actors.

[11] These assets will be used as examples throughout the paper. The derivations and mechanics of liquidity provision are the same for any pair. The most popular pool contains 'Wrapped' Ethereum (WETH) which is equivalent to ETH — transformed into an ERC20 token so that it can be easily exchanged with other ERC20 tokens. WETH and ETH are fungible. USDT is a 'stable-coin' designed to closely match the value of USD by being redeemable for US Dollars by 'Tether Corporation'. Further information on Tether and the validity of its 1:1 backing can be found in Griffin and Shams (2020).

[12] Trades with the AMM occur in discrete block-time at $t = 1, 2, 3, \ldots T$ where $t = T$ is the time at which the LP ceases liquidity provision and exits the game with the assets that can be withdrawn from the AMM at the time.

the trade in relation to the amount of x and y in the pool. This means that the trade price deviates from the ratio price as an increasing function of the proportion of the pool's liquidity that is removed.

LPs are incentivized to provide liquidity through a fixed fee of 30 basis points (bps) of the trade value, $f_t = 0.003x_t$. These fees are added to the liquidity pool, slightly increasing the pool's K with each trade. For tractability, I assume fees accrue to a separate account (which assumes LPs withdraw fees). Over the time horizons in question, this assumption makes little difference to the overall results.

Importantly, there can be many LPs to a given pool[13] and existing LPs can also increase the amount of liquidity they stake in an AMM, increasing the size of the pool. LPs provide more liquidity (referred to as "minting"), increasing the pool constant K by depositing additional assets x and y at the pool's current asset ratio of x_t and y_t. They may also withdraw their liquidity, reducing the pool's K and withdrawing x_t and y_t at the current ratio of the pool at any time. Crucially, this means they are exposed to price changes on their staked assets and thus adverse selection which I will detail further later in the paper. The daily net changes in liquidity to existing pools, described above, I refer to as pool 'flows' throughout the paper.

2.1 AMM Properties

I set out various properties for how an AMM functions for traders buying and selling ETH (y), with USD (x) being the unit of account. A derivation of each is given in Appendix B. I assume that T is a sufficiently short period such that no mints or burns occur during the period from $t = 0$ to $t = T$.

Property 1: Ignoring fees, the trade price to buy a quantity of ETH Δy_t is given as follows:

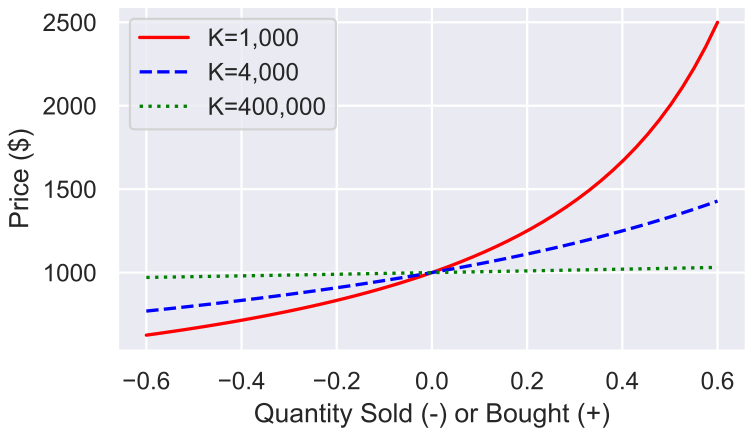
$$P(\Delta y_t) = \frac{\Delta x_t}{\Delta y_t} = \frac{x_{t-1}}{y_{t-1} - \Delta y_t} \tag{1}$$

with a trade to sell $\Delta y_t'$ units of ETH being the same but with the last term adding Δy_t. This price function gives the AMM's bonding curve, demonstrating the impact on the price of larger quantities traded.

[13] There are over 20,000 unique LP 'addresses' for the largest pools. See Wang et al. (2021) for more statistics on Uniswap LP addresses.

Figure 1: AMM Price Bonding Curve

This plot illustrates the price paid or received ($ per ETH) when buying (+) or selling (-) ETH. It assumes the initial price is $P_0 = \$1,000$. The less liquid market of $K = 1,000$ corresponds to 1 ETH and $1,000 staked, the more liquid market of $K = 4,000$ corresponds to 2 ETH and $2,000 staked and the highly liquid market of $K = 400,000$ corresponds to 20 ETH and $20,000 staked.



Property 2: The 'midpoint' price of ETH in the pool measured in USD (the price of an infinitesimally small trade that has negligible price impact) is purely a function of the quantities of the two assets in the pool at the time of the trade: $P_{0,MID} = x_0/y_0$ or more generally $P_{t,MID} = x_t/y_t$.

Property 3: Ignoring fees, the sequence in which trades occur does not matter for the final outcome (state) of the AMM — its pool quantities and midpoint price.

Property 4: Still ignoring fees, a roundtrip trade reverts both price and quantities back to their initial state, with the trader breaking even (receiving the same amount of $ as she paid).

Property 5: Two small trades in the same direction are equivalent to one larger trade in the same direction with quantity equal to the sum of the two smaller quantities (same end price in the AMM, same end state in terms of quantities in the AMM, and same cost to the trader).

Property 6: If the AMM receives a series of trades, only the buy/sell imbalance quantity of the series of trades is needed to measure the impact on the AMM's state (change in asset quantities in the pool and the pool's midpoint price), that is, the balanced part of volume (buy volume that is equal to sell volume) has no impact (irrespective of what combination of trades generates the balanced volume), the sequence of trades does not matter, and the AMM is "memoryless" in that the impact of a trade depends only on the current state of the AMM and not the history of trades.

Property 7: At every point in time, including when assets are staked and redeemed, assuming the 'midpoint' ETH price of the pool is approximately equal to the market value of ETH (that is, arbitrage has driven the AMM to the 'correct' price) the value of each of the two assets staked by the LP are equal, when measured in one unit of account, for example, the USD value of x_t is always equal to the USD value of y_t.

3 Are AMMs Gaining Traction?

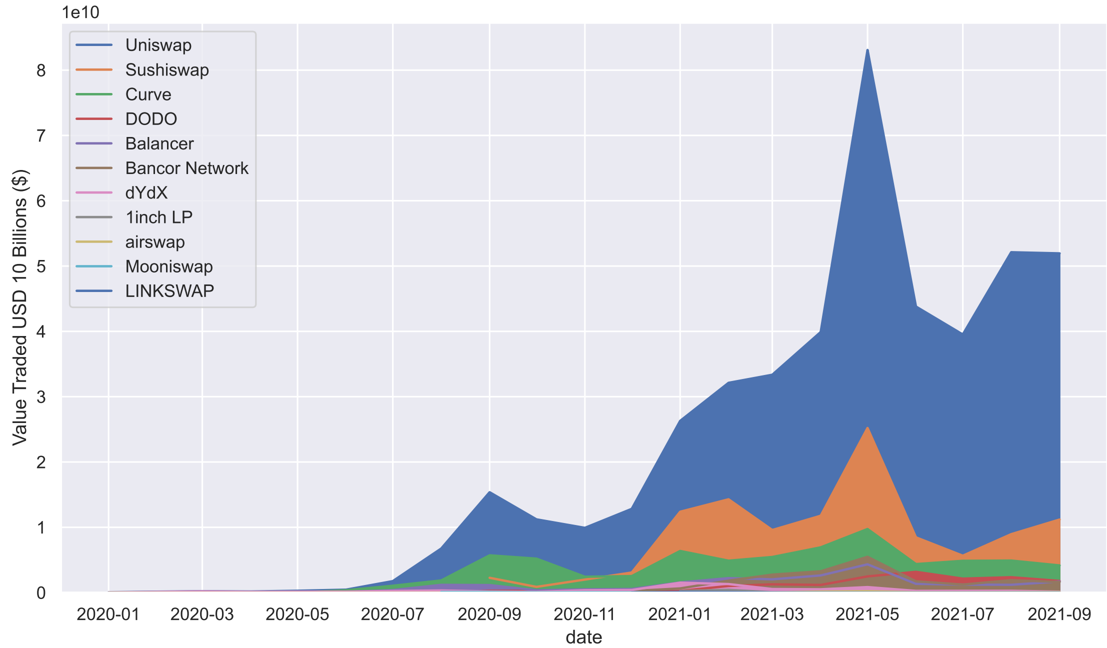
A simple measure of success is economic adoption. The total value traded via AMMs provides a useful measure of economic adoption. As shown in Figure 2, trading values have grown rapidly since the inception of AMMs in 2018, increasing from $97 million in January 2020 to $52 billion in July 2021. The value of trading in AMMs is roughly half of the total value traded on Coinbase, the largest US cryptocurrency exchange, from their most recent exchange filings.[14]

Figure 2 presents the value traded in each AMM since the 1st of January, 2020. The chart also shows that Uniswap is by far the largest AMM by value traded, averaging a 67% share of total AMM volume in 2021. While several AMMs have launched recently, only "Sushiswap", "Curve", and "Balancer" have been able to generate significant amounts of trading interest in comparison to Uniswap.[15]

[14] In their June 2021 quarterly filing, Coinbase state they executed $796 billion worth of trades in the 6 months to June 2021. Over the same period AMMs executed 399 billion. See SEC 10-Q Filing here:

[15] 'Pancakeswap' is excluded here as it is operated on the 'Binance Smart Chain' (a closed centralized system operated by Binance)by Binance themselves - a feature that has led many to question the accuracy o the volume figures quoted. In contrast, Uniswap and the AMM's listed here operate on Ethereum's public blockchain. See Aspris et al. (2021) for further institutional detail on AMMs and decentralized finance.

Figure 2: AMM Trading Values
This figure plots the total value traded on AMMs in billions $ USD each month. The sample period covers January 2020 to September 2021. The data is obtained from the Ethereum blockchain, on which each of these AMMs process transactions, via 'Dune Analytics'.



4 Can AMMs Effectively Manage Adverse Selection Risks?

For full market automation to be possible, liquidity providers must be adequately compensated for the risks they bear. The two main risks liquidity providers face in financial markets are: inventory risk and adverse selection risk. Inventory risks arise out of the necessity for an LP to hold assets to meet buying or selling demands, exposing them to changes in asset values. In modern financial markets, LPs are able to minimize inventory risks using algorithms that minimize holding times and hedge inventory exposures in correlated assets. AMMs, by contrast, require an LP to hold or provide fully collateralized positions in both assets. The fees an AMM LP earns are a linear function of the amount of inventory they hold, whereas the fees earned by a LP in traditional markets may bear little or no relationship to their inventory.

Adverse selection risks arise from information asymmetry between the liquidity provider and demanders of that liquidity. In traditional markets, the LP provides liquidity at prices based on all publicly available information experiencing adverse selection from more informed liquidity demanders with additional private information, (Glosten and Milgrom, 1985). In AMM settings, LPs are unable to set prices incorporating even public information. As a result, they bear all the costs associated with public as well as private information being impounded into prices. However, in modern markets LPs now bear a significant proportion of public information adverse selection, as shown in Aquilina et al. (2021). The economics of market making, at its core, is a problem of balancing profits from uninformed traders against the losses incurred to informed traders (adverse selection costs). Where traditional market making and liquidity provision in an AMM are similar is that they both profit from uninformed order flow — balanced (or round-trip) trading, and incur losses from informed, directional trading. Thus, all else equal, in both mechanisms, the more balanced — and the less directional — order flow (more uninformed traders and less informed traders), the more profitable is liquidity provision.

Where the two mechanisms of liquidity provision differ is that in traditional market making the profits from uninformed, balanced flow are modulated by adjusting the bid-ask spread. If the bid-ask spread is too wide, the market-maker (MM) will earn excess profits (excess to covering the adverse selection costs), attracting competing liquidity providers that undercut the MM's quotes and drive the spread to become narrower, until it hits an equilibrium level at which the MM breaks even. If the spread is too narrow, MM's make losses, leading to a widening of spreads and/or a departure of competing liquidity providers. Thus the liquidity of a stock in a traditional MM model (its bid-ask spread or the price impact of trading) is a function of the amount of informed and uninformed trading.

In AMMs, there is no explicit bid-ask spread set by the LPs, it is fixed at 30 basis points,[16] but varies in percentage terms (and is modulated by) the quantity of staked assets in the liquidity pools. This results in a slightly different equilibrating mechanism to traditional market-making, although both are driven by similar underlying principles.

If a pool is highly profitable, with fee yield exceeding the adverse selection costs, the pool should attract flows from new LPs staking the pools assets. Assuming traders' actions remain unchanged (as is traditional in models of market making for which uninformed 'noise' traders are considered exogenous, as are fundamental values and thus ultimately also

[16] This refers to Uniswap Version 2. The launch of Uniswap V3 allows for pools to set fees at 5, 30 and 100 basis points. These three options are still far from dynamic spreads in traditional markets.

informed trading), a larger quantity of staked assets will reduce the fee yield component $FY_T = \frac{F_T}{V_0}$ by increasing the denominator, V_0, without impacting the numerator. In contrast, the adverse selection component will not fall at all or as much as the pool size increases, thereby reducing LP's total profits. Conversely, if the LP is making losses in the sense that the adverse selection costs exceed the fee yield, LPs should withdraw assets from the AMM pools, thereby reducing the denominator of the fee yield and increasing the fee yield to drive up the LP profitability to a break-even level.

Thus, AMMs are expected to equilibrate the level of liquidity to match the levels of informed and uninformed trading such that LPs do not earn excess profits, but the mechanism is through modulating the pool size rather than modulating the bid-ask spread. Ultimately, the effects on liquidity are similar — if the spread is increased to cover a high adverse selection cost, this increases the cost of trading, much like decreasing the pool size to increase yield also increases price impact due to the lower liquidity constant, K. Larger price impacts of trading are economically similar to paying a higher bid-ask spread. Therefore, ultimately the relation between the level of liquidity and the underlying properties of the asset and its traders (mix of informed and uninformed, fundamental value volatility) are similar for both AMMs and traditional market-makers, despite mechanical differences in design.

4.1 Return Components for the LP in the AMM

In order to understand whether AMMs are possible, in the previous section I set out the mechanism by which AMMs can manage liquidity risks in equilibrium. In this section I derive the return components to liquidity provision so that I can apply them to the data to assess AMM feasibility.

Assume that there are no mints or burns between $t = 0$ and $t = T$. An LP that deposits assets at $t = 0$ and withdraws them at $t = T$, earns a *Total Return* of:

$$R_{TOTAL} = \frac{V_T + F_T}{V_0} - 1 \tag{2}$$

where $V_0 = x_0 + y_0 P_0 = x_0 + y_0 \frac{x_0}{y_0} = 2x_0$ is the USD value of the initial staked assets at $t = 0$, $V_T = x_T + y_T P_T = 2x_T$ is the value of the staked assets in the pool at $t = T$ (after T trades) excluding fees, and $F_T = \sum_{t=1}^{t=T} f_t = 0.003 \sum_{t=1}^{t=T} |\Delta x_t|$ is the sum of the accrued fees on the T trades. The LP's total return in (2) can be rewritten such that it is broken down into three components:

$$R_{TOTAL} = \underbrace{\left(\frac{V_{T,FIXED}}{V_0} - 1\right)}_{\text{Inventory Holding Return}} + \underbrace{\left[\left(\frac{V_T}{V_0} - 1\right) - \left(\frac{V_{T,FIXED}}{V_0} - 1\right)\right]}_{\text{Adverse Selection Cost}} + \underbrace{\frac{F_T}{V_0}}_{\text{Fee Yield}} \quad (3)$$

where $V_{T,FIXED} = x_0 + y_0 P_T$ is what the staked assets would be worth (in USD) if the LP had held them passively outside of the AMM — keeping the initial quantities fixed at x_0 and y_0.

Each of these three return components has an economic interpretation that maps to the market microstructure models of liquidity provision by market makers in traditional financial markets:

<u>The first component, the Inventory Holding Return</u>: $IHR_T = \frac{V_{T,FIXED}}{V_0} - 1$, is an increase or decrease in the value of the LP's inventory, purely from asset price changes, and not from changes in the quantities held (in contrast, adverse selection costs play out through the quantities changing in adverse ways). In traditional models of market making, the MM typically holds inventory close to zero by being allowed to short sell and buy on margin, but temporarily takes on non-zero positions while intermediating between buys and sells until such time as she can revert her inventory back to zero. While holding a non-zero inventory, the MM faces the risk that the asset price adversely appreciates or falls, which gives rise to inventory holding risk. If the MM is risk-averse, the inventory holding risk is undesirable and MM's factor it into their bid-ask spread so that they are compensated with a return for bearing this risk.[17] In contrast, in the AMM, the LP cannot hold an inventory close to zero (unless they hedge their staked liquidity positions with external contracts) as they physically have to post a positive quantity of both assets into the AMM pool. Thus, inventory holding risk is a material consideration for an AMM LP, and the IHR_T reflects the profit or loss (as a return) on the inventory the LP holds.

<u>The second component, the Adverse Selection Cost:</u> $ASC_T = (\frac{V_T}{V_0} - 1) - (\frac{V_{T,FIXED}}{V_0} - 1)$, is how much worse off the LP is from staking both assets to the AMM pool compared to the benchmark of holding the same initial asset quantities outside of the AMM. It results

[17] See classic inventory management models such as Stoll (1978), Ho and Stoll (1980), Ho and Stoll (1981), Stoll (1989), and more recently Hendershott and Menkveld (2014).

from adverse changes to the quantities of the two assets in the AMM pool as traders use the AMM to swap one asset for another. The basic idea is that when the value of the risky asset falls, yet the AMM price has not yet adjusted, arbitrageurs will buy the asset at a low price in external markets and sell it to the AMM at the AMM's incorrect (high) price, making the AMM overpay for the asset. Similarly, when the value of the risky asset increases arbitrageurs will buy the asset at the incorrect (low) price in the AMM and sell it at a higher price in an external market, resulting in the AMM selling the asset at a price below fundamental. Consequently, the quantities of the assets within the AMM change in an adverse way, resulting in an adverse selection cost akin to the classic adverse selection cost in traditional models of market making (for example, Glosten and Milgrom (1985); Kyle (1985)). As I will show, this return component is strictly less than or equal to zero, consistent with it being a true "cost" to the LP of providing traders the option to buy or sell assets from the AMM and having the traders do so on average in a way that "exploits" stale prices in the AMM. I will also show that this component is largely a function of the asset price changes — whether the asset price increases or decreases, the LP will incur a greater adverse selection cost, ASC_T.

The third component, the Fee Yield: $FY_T = \frac{F_T}{V_0}$ is expressed as a yield on the staked asset value earned by the LP. This component of total return is strictly positive as long as there are trades with the AMM. The ability to earn fee revenue is what incentivizes LPs to stake their assets knowing that they face adverse selection costs — thus, the basic proposition of an AMM is to try and earn a sufficient fee yield to at least cover the losses that will be incurred from the adverse selection costs, ideally also enough to compensate the LP for the inventory holding risk. The fee yield also maps very closely to traditional market microstructure models in which MMs earn profits from the roundtrip trades of uninformed traders and incur losses to the directional trades of informed traders. In traditional market making models, the MM adjusts the bid-ask spread that she charges per roundtrip trade so that she earns just enough from the uninformed traders (in a competitive market) to cover the losses made to the informed traders. Similarly in AMMs, roundtrip trades accrue fee yield without imposing any adverse selection costs (which are a function of price changes, and roundtrip trades, as shown above revert the price back to the original level) or inventory holding returns, whereas directional trades cause price

price changes that impose losses on the LP through the adverse selection cost component.

While the *Total Return* is the sum of each of these three components and reflects the total returns to liquidity providers, the Staking Return is the net sum of the *Adverse Selection Cost* and the *Fee Yield*. It is thus the net economic impact of the choice to invest or provide liquidity to the AMM by 'staking' assets, rather than holding them outside of the AMM. Therefore, I use it as a key measure in assessing the profitability of liquidity provision to AMMs.

Figure 3 plots *Total Return* and each of the three components against changes in price, expressed as the ratio $R = \frac{R_{t=1}}{R_{t=0}}$ of changes in price over the time period t. *Total Return* is bounded at -100% with no fee revenue and -70% with 30% assumed fee revenue in this Figure. For example, a pool holding assets ETH and USDT would have -100% returns if the price of ETH goes to zero. This is because arbitrageurs swap ETH into the pool in exchange for USDT leaving the pool holding only ETH, a worthless asset. The *Inventory Holding Return* (IHR) is bounded at -50% when the ETH price goes to zero. Notice that the IHR is larger than the *Total Return*. As the ETH price increases, arbitrageurs swap ETH out of the pool in exchange for USDT, limiting the total returns the LP receives from ETH's appreciation. *Adverse Selection Costs* is bounded at -50% because if the price of ETH goes to zero, the LP loses all the of the USDT they deposited but they would have lost the ETH regardless of investing in the pool, so just the loss in USDT is considered. At the other end of the curve, losses can exceed 100% if the price of ETH increases significantly such that the opportunity cost of holding ETH outside the pool exceeds the value of the ETH deposited.

I now vary fees at the same time as prices in Figure 4 for the two profitability measures: *Staking Return* and *Total Return*. Shades of blue represent positive returns, showing that the total return is mostly dominated by changes in price while fees are much more important for the staking return. The *Staking Return* is also more negatively impacted by positive rather than negative price changes. This means that fees are able to fully compensate for adverse selection costs for negative price changes, except at very low levels, but for price increases not even the highest levels of fees on the z-axis are able to outweigh adverse selection costs.

Figure 3: Return Components to a Theoretical LP

This figure sets out the components of returns to Uniswap liquidity provision in response to price changes in the pool, calculated as the log of the price ratio $R + 1$. The *Inventory Holding Return* is the return a liquidity provider would have obtained from holding the token assets outside of the pool, calculated as: $(R - 1)/2$. *Adverse Selection Costs* (ASC) are calculated as the Total Returns less the Inventory Holding Return, measuring returns that an LP would have otherwise obtained by not providing Uniswap liquidity. This figure assumes *Fee Yield* of 30% over the time period. *Staking Return* is the net of *ASC* and *Fee Yield*. *Total Return* is the change in the value of the pool assets.

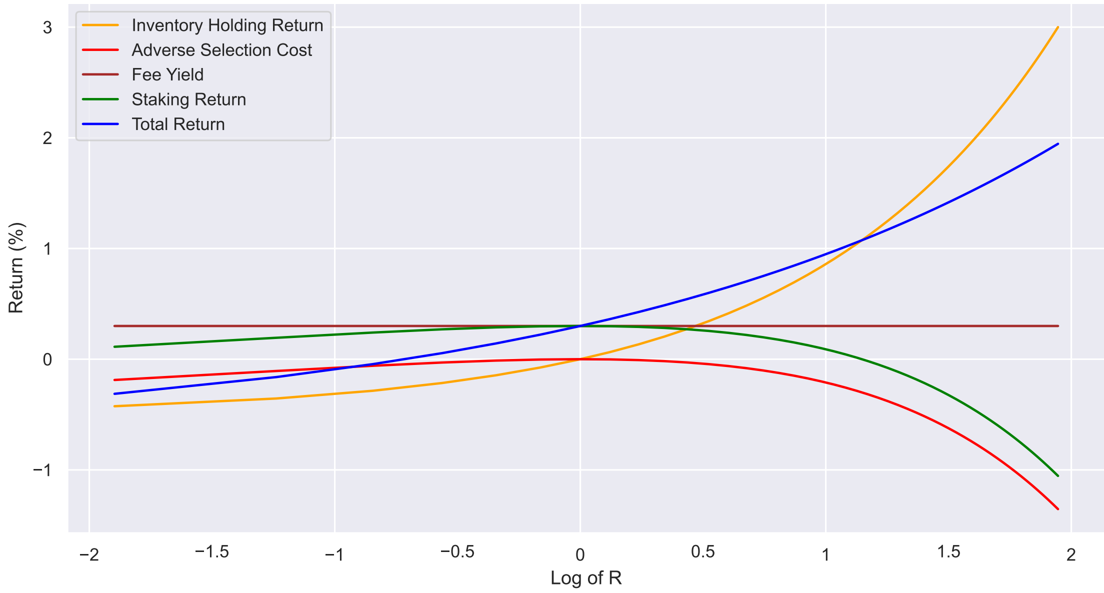
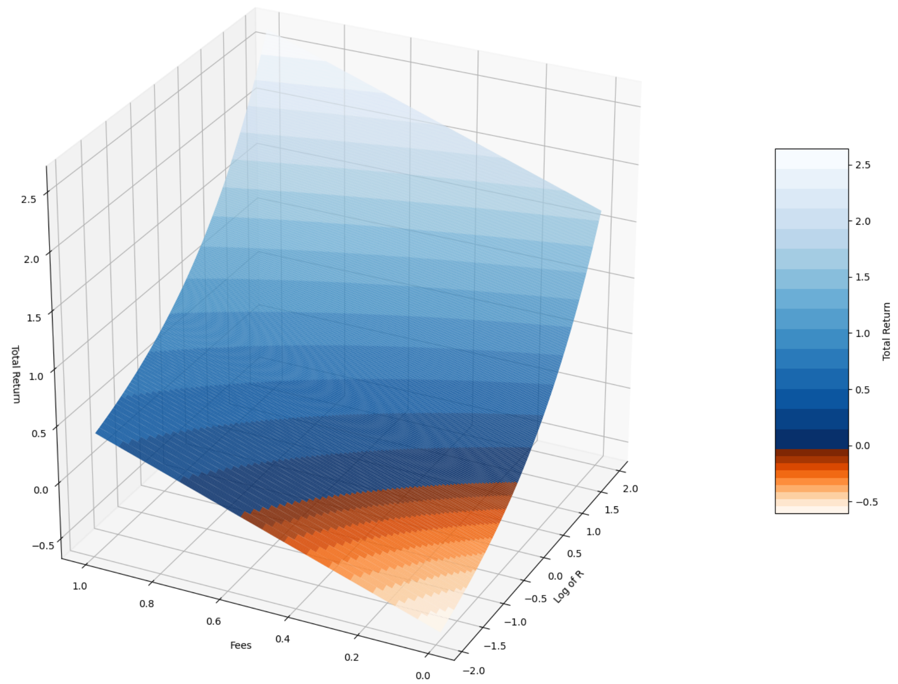
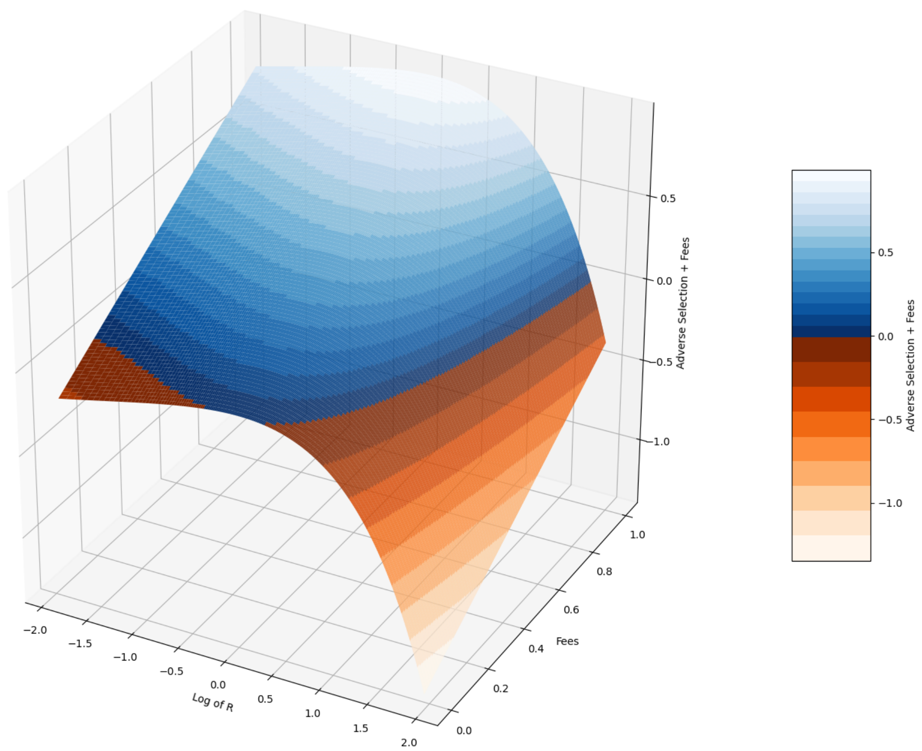


Figure 4: Return Components to a Theoretical LP With Fee Variation
This figure sets out the total return component and beneath that, the adverse selection costs plus fee returns to Uniswap liquidity provision in response to price changes in the pool. These returns are plotted on the y-axis and vary along two axes: the x-axis varies the price of assets in the pool in the form of the price ratio $R + 1$, the z-axis varies the fee revenue in the pool from 0 to 100%.

(a) Total Return



(b) Staking-Return



4.2 Adverse Selection Costs and Staking Returns in the Data

Next, I empirically assess whether AMMs sufficiently reward LPs for the significant risks they take as automated market-makers. The AMM receives little protection from the risks involved in liquidity provision. It cannot reprice stale quotes when market prices of the assets in the pool change (adverse selection risks), and it cannot actively manage inventory risks by reducing inventory holdings (inventory holding risks). The LP bears the full force of these risks with the hope that the fees they earn are adequate to offset them. In practice, this is mostly the case. The value-weighted average fee revenue, net of adverse selection costs, is positive. I refer to this measure as the 'Staking Return' as it is the net impact on returns of an LP transferring their tokens to a Uniswap pool, called 'staking' in blockchain terminology.

Table 1 sets out descriptive statistics of Uniswap pool characteristics and LP returns for the sample of the top 200 pools by mean size.[18] *Pool Size* is highly skewed, with the largest 1% of pools having more than a quarter of a billion dollars in tokenized assets, with a mean *Pool Size* of 8.5 million. LPs do not frequently update their positions, with the median pool-date experiencing no flows, with the largest 1% of *Flow* being 34% of pool value and the lowest 1% being -24%.

Fee Yield averages 12 basis points per pool-date, or 44% per year when annualized. *Adverse Selection Costs* (ASC) are always either zero or negative by construction. The pool-date mean *ASC* is -17bps per pool-date, whilst the weighted mean is -4bps (annualized -62% and -15%, respectively). This sets the sum of the two — the *Staking Return* — below 0% for the unweighted mean, but positive 2bps (7% annualized) for the pool-size weighted mean. This implies that AMM liquidity provision is profitable for an LP that invests proportionate to the size of the top 200 pools, though this profitability is not evenly distributed across pools.

Given the overall increase in cryptocurrency prices over the sample period one would expect both *Total Returns* and *Inventory Holding Returns*, which are positively related

[18] I also require the pool to have traded for more than 8-months of the 1 year sample, and remove the first 7 dates for each pool (as pool balances take time to grow to meaningful sizes). I also remove all pools that contain 'rebalancing tokens'. These are tokens that modify their balances according to predefined rules. These tokens create problems for assessing profitability. I obtain a list of rebalancing tokens from 'Coingecko'. While Uniswap V2 starts on May 19th, 2020, I remove the first few weeks of the sample and begin on the 1st of June, 2020 as pool sizes take until this time to become meaningful.

to the returns in the underlying assets, to be highly positive. The mean pool-date return is 6 and 11 basis points, but the median is negative. The majority of pools in the sample are 'Token-Token' pools, however, which means that price changes can be driven by changes in pricing relationships between two tokens, rather than a general increase in overall cryptocurrency prices. Orderflow is highly balanced, with a weighted mean *Uninformed Orderflow* of 92%. This highly balanced orderflow perhaps helps explain the viability of AMMs. As a comparison, the equivalent measure for the S&P500 Futures contract is 77.43% in Easley et al. (2012).

Table 1: Descriptive Statistics of Pools and Returns
This table sets out descriptive statistics pool-date observations in the sample of the top 200 pools in the same by mean size. There are 51,901 pool-date observations in the sample period from 1st of June 2020 to the 19th of May 2021. Pool-size weighted means are reported alongside means, due to the highly positively skewed distribution of pool size in the sample. I report 1%, Median and 99% distribution cutoffs for each variable as well as % of pool-date observations that are above 0. *Pool Size*, V_0, is calculated as the value of the two pool assets in USD at the end of the day. *Swap Value*, Q_t, is the total daily gross swap value in USD. Conversions to USD use 'Coinmarketcap' closing prices. *Flow* is the total net mint and burn amounts divided by V_0. *Total Return* is calculated as in Formula 7, *IHR* (Inventory Holding Return) as in 4, *ASC* (Adverse Selection Cost) as in 5, *Fee Yield* as in 6, *Staking Return* is the net sum of Fee Yield and ASC, *RBAL* (proportion of balanced orderflow) is $BAL_T = 2 * min[BUYS_T, SELLS_T]$ divided by the pool value V_0. *ROIB* (proportion of order imbalance) is calculated as $ROIB_T = OIB_T/V_0$ where OIB_T is the cumulative net orderflow $\sum_{t=1}^{t=T} \Delta x_t$. Return R is $\frac{P_t}{P_0} - 1$, *Turnover* is Q_t/V_0. *Volatility* is calculated as the standard deviation of returns at the rolling horizon of 30 days prior to each pool-date observation. *Toxic Orderflow* is OIB_T on Q_t and *Uninformed Orderflow* is BAL_T on Q_t.

	Mean	Wtd. Mean	Stddev	1%	50%	99%	% >0
Pool Size ($ Millions)	8.56		46.46	0.01	0.71	251.42	100.00
Swap Value ($ Millions)	1.90	44.31	16.36	0.00	0.11	43.59	100.00
Flow (% of Pool Size)	3.09	-0.08	343.62	-23.69	0.00	33.84	29.62
Total Return (%)	0.06	0.16	7.67	-11.92	-0.37	19.16	43.42
IHR (%)	0.11	0.14	6.76	-11.34	-0.42	20.48	42.55
ASC (%)	-0.17	-0.04	1.71	-2.05	-0.02	0.00	0.00
Fee Yield (%)	0.12	0.06	4.87	0.00	0.05	0.77	100.00
Staking Return (%)	-0.05	0.02	5.11	-1.74	0.01	0.48	59.76
RBAL (%)	35.56	21.28	1,146.55	0.00	13.81	253.52	100.00
ROIB (%)	-1.85	0.08	477.95	-6.00	-0.18	9.81	43.45
R (%)	0.22	0.27	13.51	-22.68	-0.83	40.96	42.55
Turnover (%)	39.61	22.24	1,621.89	0.16	15.31	260.34	100.00
Volatility (Std Deviations)	10.18	5.40	5.61	0.28	9.20	33.46	100.00
Toxic Orderflow (%)	16.92	8.35	23.73	0.09	7.24	100.00	100.00
Uninformed Orderflow (%)	83.08	91.65	23.73	0.00	92.76	99.91	100.00

Figure 5, plots the return components against the daily change in price for each pool-date in the sample. The *ASC* and *IHR* curves match those in Figure 3, varying only with R on the x-axis. *Total Return* and *Staking Return* vary with R but also upwards along the y-axis when fee-yield is non-zero. This is in contrast to the Figure 3, where a fixed fee of 30% is assumed.

Figure 5: Scatterplot of LP Return Components Against Price Ratio
This figure plots *Adverse Selection Costs*, *Inventory Holding Return*, *Staking Return* and *Total Return*, by pool-date for the top 200 pools. The ratio, $R = P_{t=1}/P_{t=0}$ for date t. *ASC* and *IHR* vary with R whilst *Total Return* and *Staking Return* vary with R and fee yield.

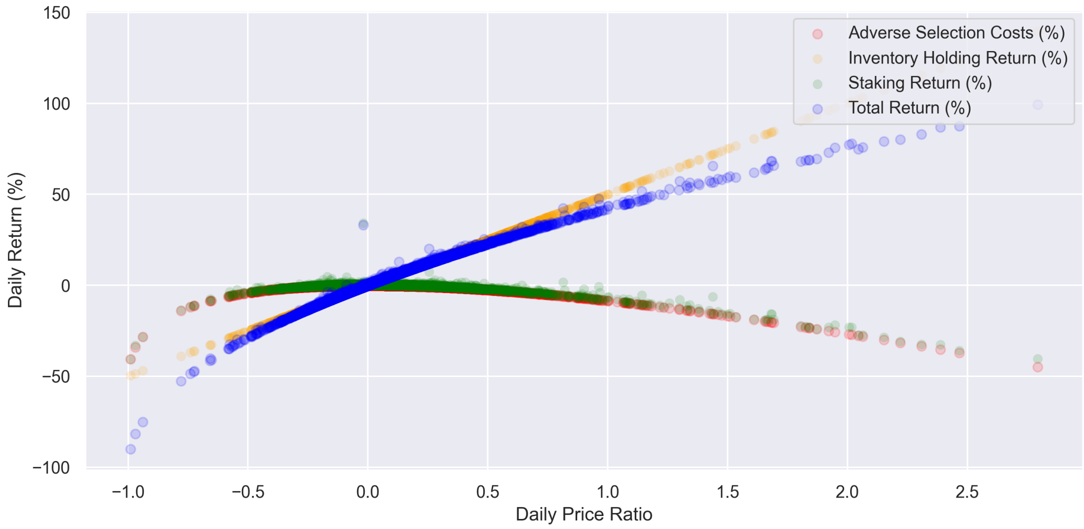


Figure 6 presents a histogram of ASC_T and $StakingReturn_T$ which is $ASC_T + FeeYield_T$ for each pool-date in the sample. *ASC* is always negative or zero, whilst the addition of *Fee Yield* shifts the distribution to the right to outweigh *ASC* to be positive for 58.8% of pool-date observations. This demonstrates that the fees earned by LPs are sufficiently large to compensate LPs for adverse selection costs they bear, for most pool-dates in the sample. It also provides evidence that liquidity provision in Uniswap AMMs may approach a rational equilibrium. Small changes in pool size (while maintaining constant orderflow) could be sufficient to generate positive average returns to LPs.

Figure 6: Histogram of Adverse Selection Costs and Staking Return for Each Pool-Date

Winsorized at 10% and 99% cutoffs. Pool-date observations. 58.86% of observations of the Staking Return are above zero.

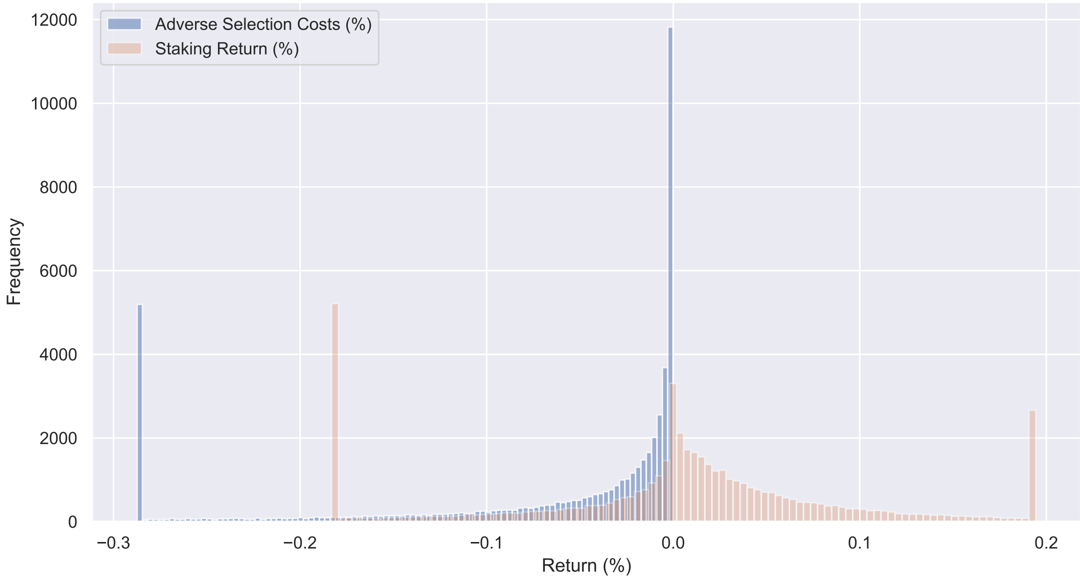


Figure 7 demonstrates the large heterogeneity in *Adverse Selection Costs* and *Staking Returns* across individual pools. Panel A shows that 22 of the largest 50 pools have an average staking return that is positive. The cumulative daily returns show that 27 pools have positive cumulative returns. The most profitable pool, 'WETH-USDT' has a cumulative staking return of 35% over the 1-year sample. A similar pool, 'USDC-WETH', which has the same tokenized asset WETH and a slightly different stablecoin (USDC rather than USDT) exhibits the same level of adverse selection costs, as would be expected, but is less profitable. This may be caused by less trading interest and/or a larger pool size in proportion to overall trading volume. 'Stable-Stable' pools such as 'USDC-USDT' and 'DAI-USDT' exhibit no significant *Adverse Selection Costs*, as expected given that the value of both assets is pegged to the USD.

Panel B of Figure 8 reports pool-size weighted average *Adverse Selection Costs* and *Staking Return* across various pool characteristics: pool-type, *Volatility* and *Pool Size*. 'Token-Stable' pools have the highest returns, averaging 4.6bps per day, due to the large *Fee Yield* of 8.2bps. 'Stable-Stable' pools return 3bps per day, almost all of which is fee yield due to the lack of adverse selection costs associated with two stable USD pegged tokens. 'Token-Token' pools experience the most *Adverse Selection Costs*, at 5.7bps per day with insufficient fees (4.9bps) to recoup these losses. While it is possible for 'Token-Token' pools to be profitable, 'SUSHI-WETH' has the second highest total cumulative return in Figure 8, the large number of token-token pools in the sample may mean that there is perhaps a bias towards higher *Volatility* tokens in this category.

Higher *Volatility* pools are less profitable, with only the lowest quartile of pools in Panel B being profitable. While this may seem intuitive, it is possible for higher volatility pools to be profitable: negative profits could drive outflows from the pool, reducing *Pool Size* and equilibrating *Fee Yields*. However, if the pool size decreases such that it becomes prohibitively expensive to execute due to the AMM bonding curve slope, volumes may also decrease so that *Fee Yield* does not increase. This might explain the positive relationship between pool size and profitability observed in Panel C. Only the largest quartile of pools are profitable.

Figure 7: Adverse Selection Costs and Fees by Pool

This figure sets out the *Adverse Selection Costs* (ASC) and *Staking Return* (ASC + Fee Yield) for the top 50 pools by average size. *ASC* and *Fee Yield* are calculated daily assuming a theoretical LP that invests at the start of the day and redeems at the end. Panel A reports value-weighted averages for each pool and Panel B reports the cumulative returns for each pool over the entire sample period of the 1st of June, 2020 to the 19th May 2021.

(a) Pool Value-Weighted Mean **(b)** Cumulative Return

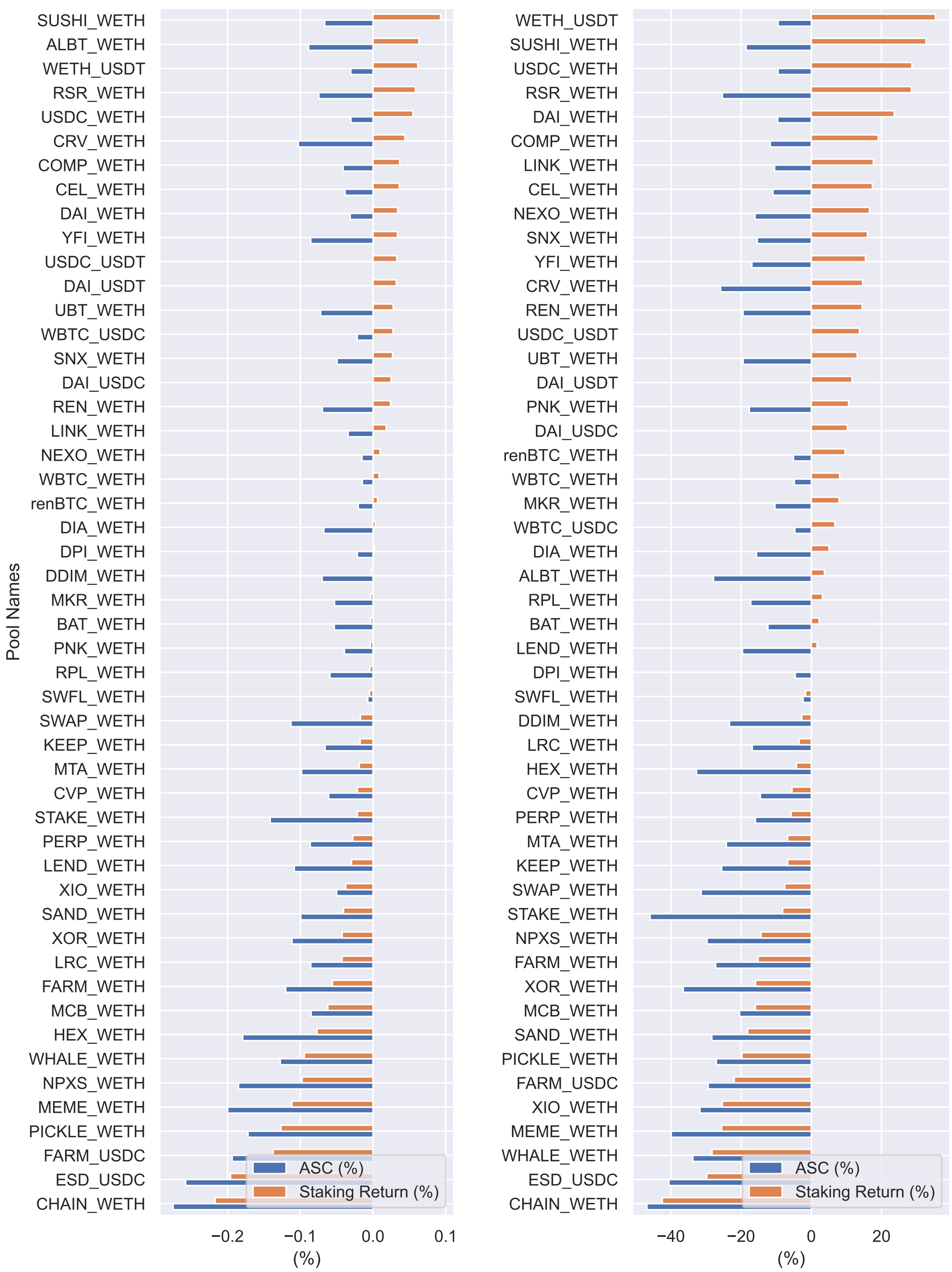
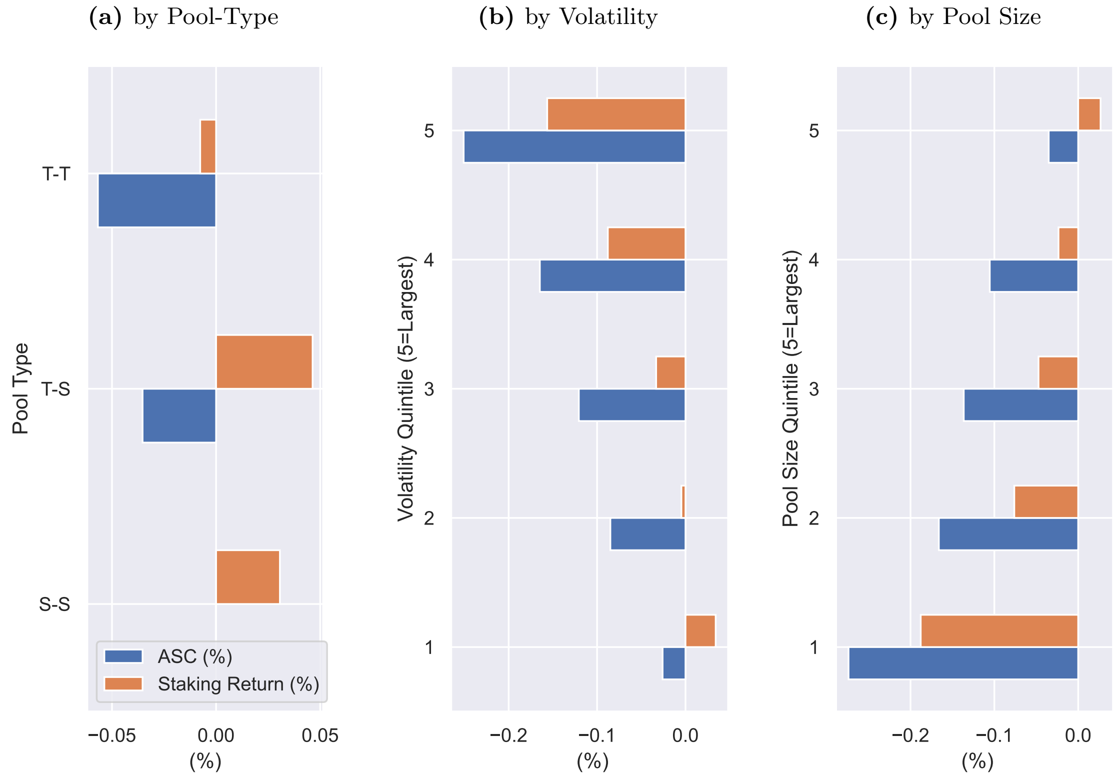


Figure 8: Adverse Selection Costs and Fees by Pool-Type, Volatility & Size
This figure sets out the *Adverse Selection Costs* (ASC) and *Staking Return* (ASC + Fee Yield) for the sample of the largest 200 pools by size. *ASC* and *Fee Yield* are calculated daily assuming a theoretical LP that invests at the start of the day and redeems at the end. Panel A reports the value-weighted average of each measure by *Pool Type*, 'T-T' refers to 'Token-Token' pools, 'T-S' refers to 'Token-Stable' pools and 'S-S' refers to 'Stable-Stable' pools. There are 181, 15 and 4 of each in the sample. Panel B reports the value-weighted average of each measure by *Volatilty Quintile*. Quintiles are constructed by calculating the rolling 30-day standard deviation of returns for each pair over the sample, and then ranking pair means into quintiles.



4.3 Deriving Determinants of LP Returns

Liquidity providers to AMMs receive *Total Returns* that can be decomposed into the three components: *IHR*, *ASC* and *Fee Yield*. These components are either driven by price changes in the underlying pool assets, the volatility of these assets, the nature of order flow in the AMM or the amount of liquidity staked in the AMM. I present theoretical relationships for each component before testing them empirically. For full derivations see Appendix D.

The IHR can be expressed as a function of returns:

$$IHR_T = \frac{1}{2}\left(\frac{P_T}{P_0} - 1\right) = \frac{1}{2}r_y \tag{4}$$

and as returns are a function of unbalanced orderflow, it can be expressed as a function of the Relative Order Imbalance ($ROIB$), where $ROIB_T = OIB_T/V_0$. OIB_T is the cumulative net orderflow $\sum_{t=1}^{t=T} \Delta x_t$ and V_0 is the pool value:

$$IHR_T = 2ROIB_T + 2ROIB_T^2 \tag{4B}$$

so that the inventory holding return is an increasing function of the order imbalance, expressed as a proportion of the total pool value in USD.

Adverse Selection Costs can also be expressed as a function of returns,

$$ASC_T = \sqrt{R_T} - \frac{1}{2}\left(R_T + 1\right) \leq 0 \tag{5}$$

where $R_T = \frac{P_T}{P_0}$, which shows that ASC are always less than or equal to zero, are minimized when prices do not change, and increase from returns in either direction (positive or negative returns). ASC_T can also be expressed as a function of the order imbalance.

$$ASC_T = -2\left(\frac{\Delta x_T}{V_0}\right)^2 = -2\left(ROIB_T\right)^2 \leq 0 \tag{5B}$$

Which shows that balanced or roundtrip trades do not contribute to ASC_T, when $ROIB_T = 0$ and $ASC_T = 0$ and when the order imbalance increases in either direction, it increases ASC_T at a quadratic rate.

The *Fee Yield* is unrelated to returns and is instead a function of the total traded volume Q_T, irrespective of both the trade direction and whether the flow is balanced or unbalanced:

$$FY_T = \frac{F_T}{V_0} = \frac{0.003\sum_{t=1}^{t=T}|\Delta x_t|}{V_0} = \frac{0.003Q_T}{V_0} = 0.003TURN_T \tag{6}$$

Thus, higher traded volumes are expected to increase *Fee Yield*. Also, lower pool value (lower V_0) will increase the *Fee Yield*. So, ultimately, *Fee Yield* is an increasing function of trading volume normalized by pool size, which I refer to as pool turnover, $TURN_T = Q_T/V_0$.

The three LP return components can be re-combined to examine the drivers of *Total Returns*:

In terms of price changes:

$$R_{TOTAL} = \sqrt{\frac{P_T}{P_0}} - 1 + 0.003\frac{Q_T}{V_0} \tag{7}$$

In terms of order flow:

$$R_{TOTAL} = 2ROIB_T + 0.003\,|ROIB_T| + 0.003RBAL_T \tag{8}$$

Interestingly, while the adverse selection costs are increasing in the absolute amount of order imbalance in either direction (excess buying or excess selling), the *Total Return* is directionally increasing in the volume of buys (decreasing in the volume of sells) due to the strong effects of the *IHR* component (Formula 8). The same effect can be seen in terms of returns: while the *ASC* increases irrespective of the direction of returns as long as the asset price changes, the total return is positively related to the return to the risky asset (Formula 7).

4.4 Applying the return framework to the data

I now apply the return relationships derived above to returns in the AMM data by testing a simple set of predictions. This will empirically confirm that AMM returns, and their components, vary mechanically in response to orderflow and prices. Later I will also examine less mechanical factors that drive returns such as pool size and volatility.

Hypotheses 1: For a given pool size (holding pool size fixed), the *Total Returns* of the LP in an AMM should be:

(A) Increasing in the volume of roundtrip trades (uninformed trading) due to increasing *Fee Yield*;

(B) Increasing in the signed order imbalance more than the absolute value of the order imbalance (informed trading) due to *Inventory Holding Returns* outweighing the *ASC* associated with directional trading;

(C) Increasing in the return of the risky asset due to inventory holding returns outweighing the ASC;

(D) Increasing in the total trading volume in the AMM.

Table 2 sets out the results of these predictions. For tractability, the derivations above assume there are no mints or burns over the daily time interval I use.[19] They also assume that fees are not capitalized into the pool value over the course of the day. So in testing these relationships I remove pool-date observations where the pool size changes by more than 52% from LP redemptions or inflows (burns or mints) which corresponds to the top 1% of the sample. I do not adjust for fees, but note the impacts of unreported regressions where extreme fee revenue pool-dates are removed.

R_{TOTAL} is constructed from daily returns following Formula 7 and is expressed in percentage terms. I first regress R_{TOTAL} against two different orderflow measures, balanced and unbalance orderflow. AMMs differ from centralized markets in that prices only move in response to unbalanced order flow, whereas on an exchange MMs can also change the prices of their resting orders. I first regress total returns on balanced orderflow, which is expected to increase total returns via the fees accrued from roundtrip trades. This positive relationship is confirmed in the data, though later regressions will show that the order imbalance has stronger explanatory power with returns. Fees, while important, accrue gradually over time and are dominated by the impact of $ROIB$ and R.

The directional order imbalance is increasing with total return as the order imbalance has a mechanical impact on prices of the assets in the AMM. This change in price then impacts *IHR* which dominates *ASC*, as predicted, such that the signed $ROIB$ is larger than the coefficient of its unsigned counterpart. The coefficient of 1.99 against $ROIB$ when regressed alone closely matches the predicted coefficient of 2 in Formula 8.[20] When $ROIB$ is also expressed as an absolute value, the positive relationship remains, but is misspecified in comparison to $ROIB$ with a much lower R^2.

A positive relationship between R_{TOTAL} and the returns of the pool assets, R, confirms that as with $ROIB$ the effects of IHR outweigh ASC. This means that, for example, an

[19]This also assumes an average LP Investment Horizon of a day. I present holding times that support this decision in Appendix F.

[20]This relationship converges to 2.00 when the top 5% of pool-dates by turnover is removed rather than the top 1%.

LP in a pool holding WETH-USDT receives returns from the appreciation in the value of WETH (the *Inventory Holding Return IHR*) but also incurs *Adverse Selection Costs* in the form of opportunity costs from not having instead fully realized those returns outside of the pool.

When expressed as a square root, R^2 the parameter estimate of 1.01 confirms that the relationship between *Total Returns* and the price of the assets is quadratic, as in Formula 7. Finally, the last column demonstrates that the *Fee Yield* is the second most important component of *Total Returns*, to the value of 30 basis points of the pool's turnover — in line with Formula 8.

Table 2: Determinants of Total Returns of LPs
This table presents regression results of pool-date AMM total returns for the sample of the top 200 pools by mean value. The dependent variable, expressed in percent: *Total Return* is calculated as in Formula 7, *RBAL* (proportion of balanced orderflow) is $BAL_T = 2 * min[BUYS_T, SELLS_T]$ divided by the pool value V_0. *ROIB* (proportion of order imbalance) is calculated as $ROIB_T = OIB_T/V_0$ where OIB_T is the cumulative net orderflow $\sum_{t=1}^{t=T} \Delta x_t$. Return R is $\frac{P_t}{P_0} - 1$, *Turnover* is Q_t/V_0. Pool-date observations with absolute value of *Flows* above the 99th percentile are removed. Pool and date fixed effects are used in each regression. Standard errors are adjusted for clustering in pool and date.

	(1) TotalRet	(2) TotalRet	(3) TotalRet	(4) TotalRet	(5) TotalRet	(6) TotalRet	(7) TotalRet
$RBAL$	3.09***						
	(19.50)						
$ROIB$		199.31***					
		(649.49)					
$\lvert ROIB \rvert$			122.84***				
			(17.40)				
R				42.86***			
				(18.83)			
R^2					101.39***		100.013***
					(288.95)		(24,416.66)
$Turnover$						3.11***	0.292***
						(18.78)	(282.85)
Constant	-0.89***	-0.00***	-2.07***	-0.06***	0.09***	-0.95***	0.001***
	(-19.77)	(-2.96)	(-18.43)	(-16.03)	(88.09)	(-19.23)	(4.48)
Obs.	51,284	51,284	51,284	51,284	51,284	51,284	51,284
R^2	0.151	0.997	0.235	0.956	0.999	0.158	1.000
Pools	200	200	200	200	200	200	200
Pool FE:	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE:	Yes	Yes	Yes	Yes	Yes	Yes	Yes

SE clustered by pool and date
T-statistics in parentheses. *p<.05; **p<.01; ***p<.001

HYPOTHESES 2: For a given pool size (holding pool size fixed), the theoretical relations above imply impacts on specific return components of the LP in an AMM:

(A) *Fee Yield* should be increasing in the volume of roundtrip trades (uninformed trading), the absolute value of order imbalance, and the total dollar volume of trading;

(B) *Adverse Selection Costs* should be decreasing (becoming larger negative costs) quadratically in the absolute dollar volume of directional (unbalanced) order flow (informed trading);

(C) *Inventory Holding Returns* should be increasing in the (directional) order imbalance, increasing in the squared order imbalance.

Table 3 empirically tests these predictions. The first three columns regress *Fee Yield* on its theoretical components. As predicted, *Fee Yield* is increasing in the proportion of balanced trading, the absolute value of the order imbalance and turnover.

Adverse Selection Costs are regressed on two difference specifications of the orderflow imbalance, $ROIB$, in Columns 4 and 5. While both forms have negative relationships, as predicted, the squared form of $ROIB$ has more explanatory power, providing evidence that the relationship is indeed quadratic. The point estimate of the $ROIB^2$ coefficient is close to the value of 2 predicted in Formula 5B.[21]

The last four regressions examine the *Inventory Holding Return*. IHR has a positive relationship with $ROIB$ and $ROIB^2$ but the combination of the two, $2ROIB + 2ROIB^2$ as in Formula 4B, returns the predicted coefficient of nearly 1 (exactly 1 when extreme fee pool-date observations are excluded).

Therefore, the AMM return relationships derived in Formulas 4 to 8 perform well when applied to the data. They accurately predict actual returns over daily time horizons. The simplifying assumptions of zero mint and burns and no fee capitalization create minor deviations, but these deviations become insignificant once extreme mint, burn and fee revenue pool-dates are excluded.

[21] The coefficient becomes exactly 2 once extreme fee yield pool-dates are excluded from the regression of pool-dates. This is because the reinvestment of significant fee yield within the 1-day horizon violates the derivation assumptions.

Table 3: Determinants of Return Components of LPs
This table presents regression results of pool-date AMM total returns for the sample of the top 200 pools by mean value. The dependent variables are the return components expressed in percent: *Fee Yield* is calculated as in Formula 6, *ASC* (Adverse Selection Cost) as in 5, *IHR* (Inventory Holding Return) as in 4. The independent variables: *RBAL* (proportion of balanced orderflow) is $BAL_T = 2*min[BUYS_T, SELLS_T]$ divided by the pool value V_0. *Turnover* is Q_t/V_0. *ROIB* (proportion of order imbalance) is calculated as $ROIB_T = OIB_T/V_0$ where OIB_T is the cumulative net orderflow $\sum_{t=1}^{t=T} \Delta x_t$, $|ROIB|$ is the absolute value of ROIB, $ROIB^2$ is ROIB squared, $2ROIB + 2ROIB^2$ is as in Formula 4B. Pool-date observations with absolute value of *Flows* above the 99th percentile are removed. Pool and date fixed effects are used in each regression. Standard errors are adjusted for clustering in pool and date.

	(1) Fee Yield	(2) Fee Yield	(3) Fee Yield	(4) ASC	(5) ASC	(6) IHR	(7) IHR	(8) IHR
$RBAL$	0.30***							
	(267.40)							
$\|ROIB\|$		5.01***		-48.55***				
		(4.83)		(-4.25)				
$Turnover$			0.293***					
			(309.47)					
$ROIB^2$					-165.57***		403.53***	
					(-21.49)		(4.72)	
$ROIB$						222.59***		
						(29.97)		
$2ROIB + 2ROIB^2$								95.02***
								(41.65)
Constant	0.00***	0.01	0.001***	0.66***	-0.02***	0.07***	-0.25***	-0.09***
	(16.20)	(0.42)	(4.52)	(3.47)	(-3.91)	(12.44)	(-3.49)	(-27.35)
Obs.	51,284	51,284	51,284	51,284	51,284	51,284	51,284	51,284
R^2	0.997	0.280	0.998	0.440	0.984	0.956	0.375	0.993
Pools	200	200	200	200	200	200	200	200
Pool FE:	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE:	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

SE clustered by pool and date
T-statistics in parentheses. *p<.05; **p<.01; ***p<.001

I next examine evidence suggestive of an equilibrium state in the cross-section of pools.

Hypotheses 3: As a result of the equilibrating mechanisms discussed above, I should see that liquidity in AMM pools (or *Pool Size*) is:

(A) Decreasing in asset *Volatility* (a driver of adverse selection risk);

(B) Decreasing in the ratio of directional to total order flow (order flow toxicity), or equivalently, increasing in the ratio of balanced to total order flow (noise trader share).

Figure 9 uses a simple univariate analysis to show a clear negative relationship between pool size and volatility. However, in order to determine causality, a more formal analysis is required.

Figure 9: Scatterplot of Pool Size versus Pool Volatility
Volatility, on the x-axis, is calculated as the standard deviation of returns at the horizons of 30 days prior to each pool-date observation, and winsorized at the 99% cutoff of their distribution. Means are then taken for each pool. The y-axis reports the mean logged value of assets in the pool.

Table 4 reports regression results that test Hypothesis 3. As predicted, higher asset *Volatility* measured by the standard deviation in returns over the previous 30 days is associated with smaller *Pool Size*. The coefficient shows that a 10 standard deviation increase in returns is associated with a 13.9% decrease in the *Pool Size*. This is a limitation in the design of Uniswap AMMs being unable to vary the fixed 30 basis point upwards to compensate for the additional risks inherent in more volatile pools — the only adjustment possible is via the *Pool Size* (or the amount of liquidity provided). In traditional financial markets, market-makers simply charge a higher spread — this increases the price of liquidity but does not necessarily reduce the amount of liquidity available — unlike AMMs. Nonetheless, there is evidence for the equilibrating mechanism at work.

Orderflow *Toxicity* has a negative and significant relationship with pool size: an increase in *Toxicity* of 10% (from an average of 18.7%) decreases *Pool Size* by 3%. Similarly, *Uninformed Share* has a significant and positive relationship on *Pool Size*, as predicted. An increase in *Uninformed Share* of 10% (from an average of 40.6%) increases *Pool Size* by 5.1%. Performing pooled regressions of pool-means alone in the last three columns of Table 4 show much larger coefficients once fixed effects that control for cross-sectional and time-series variation are removed. Perhaps it is unsurprising that in this innovative new space, token-specific and day-specific factors represent important sources of variation. The incentives provided for Uniswap LPs to switch to a rival AMM, the '*Sushi-attack*' variable, are unsurprisingly associated with smaller pool sizes during the incentive period. Conversely, the incentives provided to Uniswap LPs, the '*Uni-Incentive*' variable, are associated with larger *Pool Size*. Han et al. (2021) finds the same result.

Table 4: Impact of Pool Attributes on Pool Size

This table presents results for tests of Hypothesis 6. The dependent variable, *Pool Size*, is the natural log of the value of the pool in billions on a given date. *'SushiAttack'* takes the value of 1 for the period of time where AMM rival Sushiswap provided incentives to Uniswap LP's to switch platforms. *'UniIncentive'* takes the value 1 for period of incentives provided by Uniswap. *Volatility* is measured as the standard deviation of returns at the horizons of 30 days prior to each pool-date observation, divided by 10, and winsorized at the 99% cutoff of their distribution. *Toxicity* is calculated as the daily order imbalance as a proportion of the total order flow, $\frac{OIB}{Q}/10$. *Uninformed Share* is calculated as the daily balanced orderflow as a proportion of the total order flow, $\frac{BAL}{Q}/10$. Pool-date observations with absolute value of *Flows* above the 99th percentile are removed. The regression includes pool and date fixed effects in the first three columns. The last three columns are pooled regressions of pool-means with no fixed effects. Standard errors are adjusted for pool clustering.

	(1) Pool Size	(2) Pool Size	(3) Pool Size	(4) Pool Size	(5) Pool Size	(6) Pool Size
$SushiAttack$	-0.62***	-0.62***	-0.62***			
	(-3.31)	(-3.32)	(-3.32)			
$UniIncentive$	1.63***	1.68***	1.68***			
	(27.89)	(28.45)	(28.45)			
$Volatility$	-0.07***			-1.98***		
	(-4.92)			(-7.35)		
$Toxicity_{t-1}$		-0.05***			-0.42***	
		(-16.78)			(-6.05)	
$UninformedShare_{t-1}$			0.05***			0.42***
			(16.78)			(6.05)
Constant	-7.09***	-7.08***	-7.57***	-4.68***	-6.00***	-10.22***
	(-474.82)	(-1,045.32)	(-302.12)	(-14.28)	(-34.96)	(-18.15)
Obs.	51,284	51,269	51,269	200	200	200
R^2	0.009	0.015	0.015	0.259	0.156	0.156
Pools	200	200	200			
Pool FE:	Yes	Yes	Yes	No	No	No
Date FE:	Yes	Yes	Yes	No	No	No

SE clustered by pool and date

T-statistics in parentheses. *p<.05; **p<.01; ***p<.001

I now examine hypotheses that aim to test for the presence of equilibrium effects in the other direction, via the pool return components. Higher fee yielding pools should attract flows, becoming larger sized pools.

Hypothesis 4: For a given trading volume and order imbalance (holding various measures of trading activity fixed), the *Fee Yield* in an AMM should be decreasing in the *Pool Size* (value of staked assets).

Table 5 regresses *Fee Yield* and other return components on *Pool Size*, controlling for the total trade volume Q and subsequently with the order imbalance OIB in billions. *Fee Yield* is found to have a negative relationship with *Pool Size*, as predicted, and this is robust to the inclusion of the order imbalance, OIB. This implies that pools with lucrative fee revenue become larger which in turn depresses the *Fee Yield*.[22] The log coefficient implies that a 1% increase in *Pool Size* decreases *Fee Yield* by 1 percent.

ASC has a positive relationship with *Pool Size*, which implies that larger pools experience less *Volatility*. This is confirmed by the robust negative relationship between *Pool Size* and *Volatility* observed in Table 4.

Further evidence of equilibrium effects are that when these two effects are considered together, there is no relationship between pool size and the net of ASC and fee-yield (the *Staking Return*). Nonetheless, for all pools to be in perfect equilibrium they should each have a *Staking Return* (or ASC + Fees) that is approximately zero.[23] Pools with positive *Staking Returns* are too small in equilibrium, and pools with negative *Staking Returns* are too large. Figure 8 shows that this is often the case for various types of pools. In particular, Token-Stable and Stable-Stable pools should become larger, while more volatile pools should decrease in size, and smaller pools should further decrease in size.

Lastly, total return and IHR both exhibit strong negative relationships with *Pool Size*. There are perhaps two types of agents that become LPs. The first, an investor that is already long the pool's underlying assets seeking to add additional yield to their position. For these agents inventory risk (in the form of *IHR*) is not as important as the *Staking Return*. The second set of investors is not already long the underlying pool assets and is considering becoming an LP as a standalone investment strategy. For these agents, the *Staking Return* is of most importance.

[22] I explore these equilibrium effects in further detail in Section 6.

[23] It is worth noting that this assumes LPs rationally assess the impact of their decision to stake in Uniswap correctly. In the context of rising tokenized asset prices, LPs may not realize their S*taking Returns* are negative if their *Total Returns* are highly positive, ie. they may not fully appreciate the opportunity costs of staking.

Table 5: Impact of Pool Size on LP Return Components
This table presents regression results of pool-date AMM return components for the sample of the top 200 pools by mean pool size. The dependent variables, expressed in percent: *Total Return* is calculated as in Formula 7, *Fee Yield* is calculated as in Formula 6, *ASC* (Adverse Selection Cost) as in 5, *IHR* (Inventory Holding Return) as in 4, *stake-ret* is the Staking Return which is the net of ASC and fee yield. The independent variables: *Pool Size* is the log of pool value in USD at the start of the day, Q is the total value of all swaps on the pool-date in billions USD. *OIB* is the total value of the cumulative net orderflow $OIB_T = \sum_{t=1}^{t=T} \Delta x_t$ in billions USD. Pool-date observations with absolute value of *Flows* above the 99th percentile are removed.

	(1) Total-Ret	(2) Total-Ret	(3) FeeYield	(4) FeeYield	(5) ASC	(6) ASC	(7) IHR	(8) IHR	(9) Stake-Ret	(10) Stake-Ret
Pool Size	-0.15***	-0.15***	-0.01***	-0.01***	0.01***	0.01***	-0.16***	-0.16***	0.00	0.00
	(-6.87)	(-7.05)	(-11.43)	(-11.44)	(2.67)	(2.68)	(-6.42)	(-6.58)	(0.99)	(1.00)
Q	3.92	6.67***	0.66**	0.67**	-1.09**	-1.15**	4.35	7.15***	-0.43	-0.49
	(1.58)	(2.69)	(2.47)	(2.50)	(-1.96)	(-2.08)	(1.58)	(2.60)	(-1.43)	(-1.62)
OIB		1,593.57***		5.51***		-39.68***		1,627.74***		-34.17***
		(18.91)		(6.87)		(-5.93)		(18.61)		(-5.49)
Constant	-1.07***	-1.10***	0.03***	0.03***	-0.07*	-0.06*	-1.03***	-1.07***	-0.04	-0.04
	(-6.77)	(-7.06)	(4.46)	(4.44)	(-1.96)	(-1.93)	(-6.01)	(-6.27)	(-1.31)	(-1.28)
Obs.	51,284	51,284	51,284	51,284	51,284	51,284	51,284	51,284	51,284	51,284
R^2	0.001	0.018	0.004	0.004	0.000	0.000	0.001	0.014	0.000	0.000
Pools	200	200	200	200	200	200	200	200	200	200
Pool FE:	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE:	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

SE clustered by pool and date
T-statistics in parentheses. *p<.05; **p<.01; ***p<.001

Table 6 regresses the return components on the volatility of the pool's returns. *Volatility* is calculated as the standard deviation of daily returns in the 30 days prior to a given pool-date. The table shows a negative relationship between ASC and volatility. This is because ASC is always negative or zero by construction, higher volatility makes ASC more negative.

The positive and significant relationship between *Volatility* and *Fee Yield* implies that LPs are compensated for bearing the larger *Adverse Selection Costs* associated with additional *Volatility* (with the negative relation implying larger costs). These relationships are both statistically and economically significant — with a mean volatility of 10.18 the coefficient of 7 basis points implies that a doubling of *Volatility* roughly doubles the expected daily *Fee Yield.* However, as the relationship with staking return is negative, the compensation LPs receive for bearing more risk is not sufficient to overcome *Adverse Selection Costs* (and/or as found in Table 5), *Pool Size* is still too large on average for realized levels of *Volatility.*

Table 6: Impact of Return Volatility on LP Return Components
This table regresses LP return components on pool-volatility, for the sample of the top 200 pools by mean pool size. The dependent variables: *Total Return* measures the total return to liquidity provision (see Equation 7), *Fee Yield* measures pool fee yield (see Equation 6), *ASC* (Adverse Selection Cost) measures the costs of staking (see Equation 5), *IHR* (Inventory Holding Return) measures the returns to holding the assets in the pool (see Equation 4), *Stake Ret* is the staking return, which is the net of *ASC* and *Fee Yield. Volatility* is measured as the standard deviation of returns at the horizons of 30 days prior to each pool-date observation, divided by 10. Pool-date observations with absolute value of *Flows* above the 99th percentile are removed. Volatility is winsorized at the 99% level.

	(1) Total Return	(2) Fee Yield	(3) ASC	(4) IHR	(5) Stake-Ret
Volatility	0.12	0.07***	-0.11***	0.16*	-0.03*
	(1.46)	(24.17)	(-5.25)	(1.69)	(-1.86)
Constant	-0.10	0.02***	-0.06**	-0.06	-0.04*
	(-1.21)	(6.64)	(-2.46)	(-0.65)	(-1.78)
Obs.	51,284	51,284	51,284	51,284	51,284
R^2	0.000	0.018	0.001	0.000	0.000
Number of pairid	200	200	200	200	200
Pool FE:	Yes	Yes	Yes	Yes	Yes
Date FE:	Yes	Yes	Yes	Yes	Yes

SE clustered by pool and date
T-statistics in parentheses. *p<.05; **p<.01; ***p<.001

This section derives and empirically verifies the return components for liquidity providers. It also presents evidence that AMM pools rationally equilibrate towards fee revenue covering adverse selection, suggesting that AMMs use the interplay between uninformed orderflow and pool size to manage adverse selection risks.

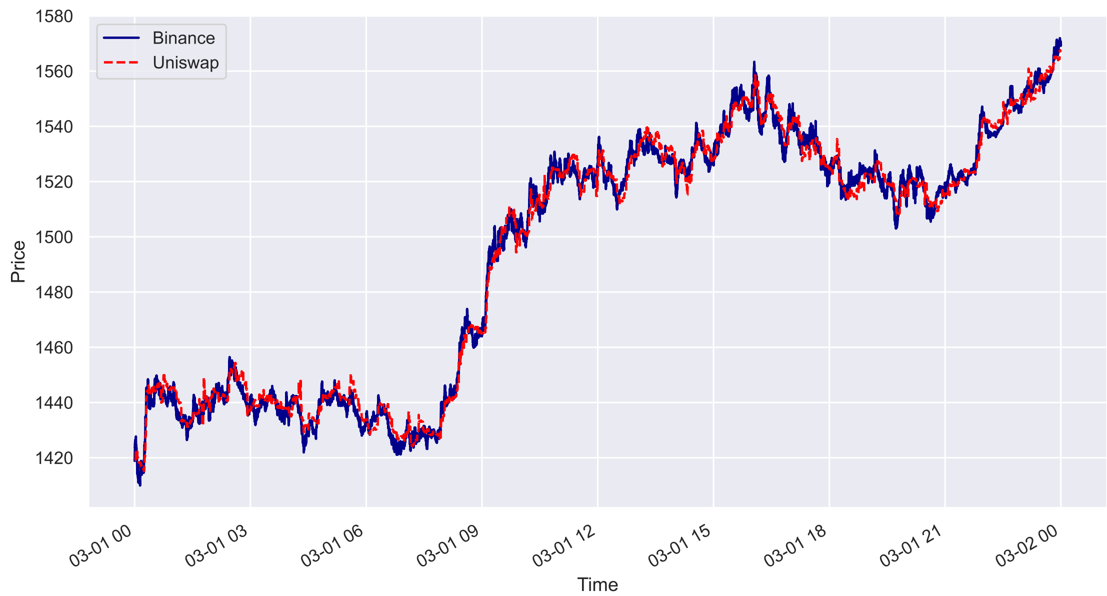
5 How Well do AMMs Determine Prices?

An important function of markets is to provide accurate prices. To assess the accuracy of prices I use the centralized market as a comparison. Figure 10 compares the Uniswap price in WETH-USDT (the largest Uniswap pool) at the end of each minute to the orderbook midpoint from Binance, the largest cryptocurrency exchange, for the 1st of March, 2021 in ETH-USDT. Prices appear to follow each other closely. Prices are calculated every minute to allow sufficient time for Uniswap arbitrage transactions to occur to equalize prices. This is because the Ethereum blockchain takes 14 seconds on average to process new transactions.

I now examine the efficiency of prices over the whole sample, calculating the proportion of the day in which profitable arbitrage opportunities exist. I calculate the difference in prices each second less any frictions or costs to arbitrage. These are the 30bps Uniswap fee, the Binance trading fee of 10bps. Another friction is the cost of 'gas'— a fee for processing a transaction on the Ethereum blockchain. This is difficult to estimate on an individual transaction basis so this friction is ignored. I then sum all actionable price differences — those that persist for longer than the maximum Ethereum block time of 19 seconds.

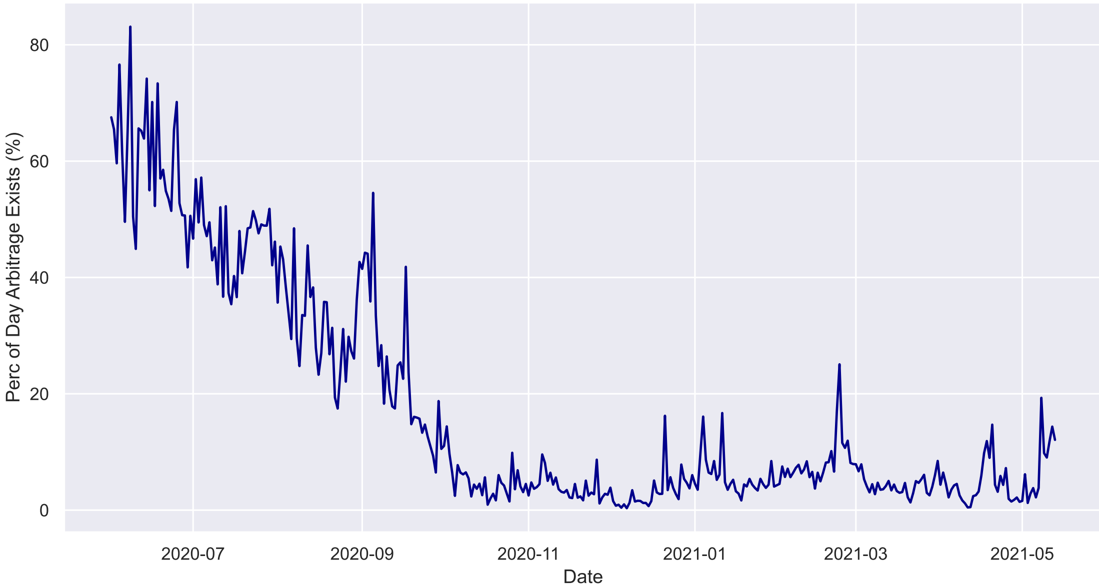
Figure 10 charts the proportion of the trading day profitable and actionable arbitrage opportunities exist between Binance and Uniswap in WETH-USDT. The Figure shows that the percentage of the trading day often reduces to single-digit percentages from October 2010 onward, coinciding with significant increases in Uniswap trading volumes and pool sizes. The average size of the arbitrages that remain are around 10 basis points[24] and it is reasonable to assume that the Ethereum gas price accounts for this remaining difference.

Figure 10: Uniswap vs Binance Prices for WETH-USDT
This figure plots the Uniswap pool price or WETH-USDT, the largest Uniswap Pool, and the Binance ETH-USDT orderbook midpoint price each minute for a 24 hour period on the 1st of March, 2021.



[24]See Appendix Figure 14 for a plot of the size of remaining actionable arbitrage opportunities over time.

Figure 11: Uniswap vs Binance Prices for WETH-USDT Daily Timeseries
This figure plots the percentage of the day in which actionable and profitable arbitrages exist between Uniswap WETH-USDT, the largest pool, and Binance ETH-USDT. Arbitrages must persist for more than 19 seconds and must be greater than 40 basis points to be profitable.



While I do not examine the extent to which price discovery occurs on Uniswap, or the extent to which it leads price changes in comparison to centralized markets, it is clear that Uniswap prices align closely with centralized markets. Prices could align even closer with the reduction in arbitrage frictions such as the Ethereum gas fee, which should become insignificant as 'proof-of-stake' is adopted, and trading fees.

6 Equilibrium Effects: Is liquidity Allocated to Where it is Needed?

In any market, there needs to be a mechanism to allocate capital to its most productive use. The sensitivity of fund inflows and outflows to fund performance is of key interest in the investment funds management literature. In Uniswap AMMs, flows are crucial in determining LP returns and pushing pools towards equilibrium. Section 4.4 showed evidence consistent with such an equilibrium. I now search for evidence of the pools tending towards equilibrium through the pool size.

HYPOTHESIS 5: Liquidity should flow into pools that have higher total returns.

To examine if flows chase returns I first compute correlations between current period pool flows and 30 lags and leads of fee yield in the time-series. Figure 12 shows that fee yield is positively correlated with lagged flows but negatively correlated with future flows. This demonstrates the equilibrating effect of flows on fee yield — flows chase yields but also act to depress them.

Figure 12: Correlogram of Current Period Flows Against 30 Day Lag/Lead Fee Yield
Pearson correlation coefficients are constructed between current period flows and the respective lag/lead Fee Yield first within a given pool. Means are then constructed across all pools.

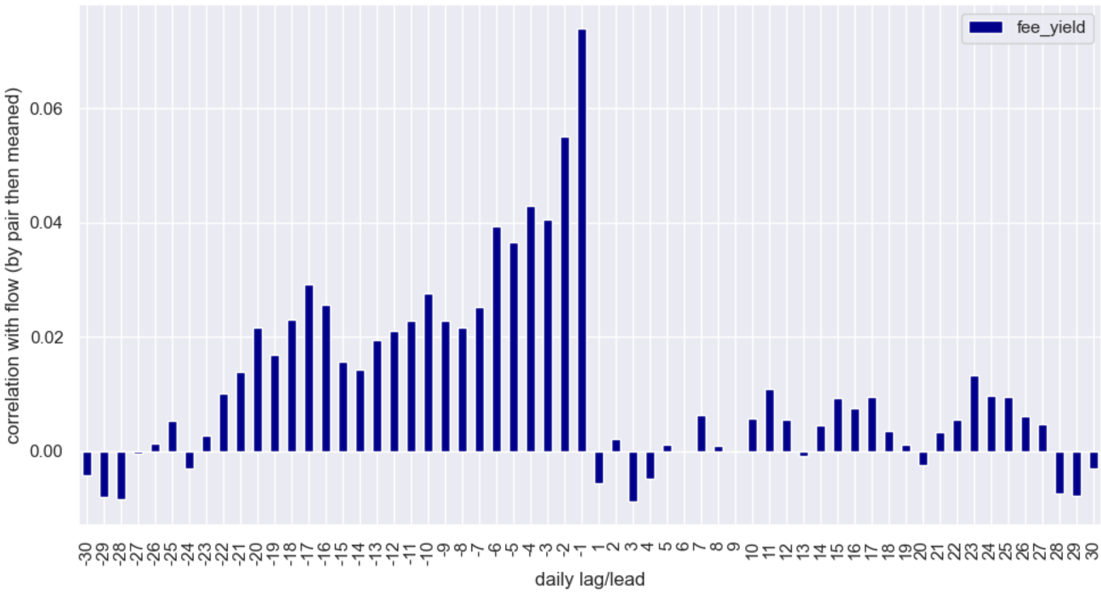


Table 7 regresses pool *Flow* on lagged measures of each pool return component, as well as lagged pool *Flow*, similar to a structural VAR model. Consistent with the correlogram, fees are positively related to flows. A one percent increase in fees in the previous day results in a 1.1% flow into the pool. However, as the mean *Fee Yield* in the sample is 12bps, there is a muted reaction by flows to such an increase, which perhaps explains why the cross-section of AMMs is not in full equilibrium, with some pools exhibiting both positive and negative *Total Returns*. Capponi and Jia (2021) also find this positive relationship between flows and fee yield in a contemporaneous model at the weekly level.

A positive relationship between *IHR* and *Total Returns* is also found, however their coefficients are not meaningfully large. *ASC* exhibits either no, or a very small, relationship with *Flow*. Table 6 documents a strong relationship between *ASC* and 30 day pool *Volatility*, which implies that LPs may make investment decisions over longer horizons rather than the daily horizon used in this specification. As *ASC* forms a key component of staking returns, perhaps the same explanation can be offered for the lack of a significant relationship between *Staking Return* and *Flow*. Fee yields are visible and easy to calculate, so perhaps provide the most salient driver of LP flows.

The regression includes dummy variables '*Sushiswap Attack*' and '*Uniswap Incentive*'. These variables take the value of one during the period a rival AMM and Uniswap provide incentives to LPs in certain pools. Date fixed effects likely subsume most of the significance of these variables, so there is no impact.

Table 7: Time Series Regression of Flows on LP Return Components
This table reports regressions of pool flows on lagged return components. *Flow* is the daily net mint and burn value divided by the pool value, in percent. The following are in basis points: *Total Return* as in Formula 7, *Fee Yield* as in Formula 6, *ASC* as in 5, *IHR* as in 4 and *StakingReturn* is the net of *ASC* and *Fee Yield*. *'SushiAttack'* equals 1 for period AMM rival Sushiswap provided LP incentives. *'UniIncentive'* equals 1 for time period of Uniswap LP Incentives. Pool-date observations with absolute value of *Flows* above the 99th percentile are removed.

	(1) Flow	(2) Flow	(3) Flow	(4) Flow	(5) Flow
SushiAttack	3.72	3.63	3.66	3.71	3.67
	(1.49)	(1.46)	(1.47)	(1.49)	(1.47)
UniIncentive	0.10	0.25	0.09	0.09	0.08
	(0.22)	(0.55)	(0.19)	(0.21)	(0.19)
$Flow_{t-1}$	0.04***	0.03***	0.04***	0.04***	0.04***
	(3.67)	(3.39)	(3.76)	(3.70)	(3.76)
$Flow_{t-2}$	0.03***	0.03***	0.03***	0.03***	0.03***
	(3.48)	(3.33)	(3.54)	(3.50)	(3.53)
$Flow_{t-3}$	0.02***	0.02***	0.02***	0.02***	0.02***
	(3.15)	(3.04)	(3.15)	(3.16)	(3.15)
$TotalReturn_{t-1}$	0.01***				
	(3.13)				
$TotalReturn_{t-2}$	0.01***				
	(3.79)				
$TotalReturn_{t-3}$	0.01*				
	(1.82)				
$FeeYield_{t-1}$		1.06***			
		(2.62)			
$FeeYield_{t-2}$		0.36			
		(1.44)			
$FeeYield_{t-3}$		0.28			
		(1.44)			
ASC_{t-1}			-0.03*		
			(-1.69)		
ASC_{t-2}			-0.02		
			(-1.39)		
ASC_{t-3}			-0.01		
			(-0.55)		
IHR_{t-1}				0.01***	
				(2.86)	
IHR_{t-2}				0.01***	
				(3.37)	
IHR_{t-3}				0.01	
				(1.53)	
$StakingReturn_{t-1}$					-0.69
					(-0.62)
$StakingReturn_{t-2}$					-1.28
					(-0.97)
$StakingReturn_{t-3}$					0.59
					(0.56)
Constant	-0.00	-0.16***	-0.01	-0.00	-0.00
	(-0.09)	(-4.54)	(-0.45)	(-0.17)	(-0.10)
Obs.	50,019	50,019	50,019	50,019	50,019
R^2	0.004	0.007	0.004	0.004	0.004
Number of pairid	200	200	200	200	200
Pool FE:	Yes	Yes	Yes	Yes	Yes
Date FE:	Yes	Yes	Yes	Yes	Yes

SE clustered by pool and date
T-statistics in parentheses. *p<.05; **p<.01; ***p<.001

While positive returns may attract inflows (and negative returns outflows), it is also possible that staking flows by LPs themselves impact pool returns. This occurs as the explicit fees charged on trading (the primary revenue of the LPs) are independent of pool size. Larger pools spread these revenues out amongst a wider base of LPs, reducing overall pool return. To test for these effects, I propose the following:

HYPOTHESIS 6: Following liquidity inflows (/outflows) to/from an AMM pool, the *Total Return* of the LPs in that pool should be lower (/higher), as should be *Fee Yield*.

Table 8 regresses each return component on lagged flows, as well as lagged pool returns, similar to a structural VAR model. *Fee Yield* equilibrates as predicted. An increase in pool *Flow* of 10% in the previous day results in a decrease in the daily *Fee Yield* of 0.6 basis points. As the weighted mean *Fee Yield* averages 6 basis points, this represents a similar reduction of 10 percent.

Perhaps unsurprisingly, the external incentives provided by Uniswap to encourage liquidity provision are associated with a reduction in *Fee Yield* and *Staking Return*. This is because LPs received the newly created 'Uniswap Token' during this period and were thus willing to accept lower pool yields in return for these additional incentives. The '*Sushi Attack*' period has the opposite impact, as LPs were incentivized to leave Uniswap to instead stake on their competitor, SushiSwap, during this period. While they also impact *Total Returns* and *IHR*, this is likely confounded by price changes to major assets such as Ethereum during the period.

In summary, this section provides evidence that pool size works to equilibrate fee yields through inflows and outflows. LPs direct flows to high yield pools which then dilutes down the percentage *Fee Yield* received by all LPs in the pool.

Table 8: Time Series Regression of LP Return Components on Flows
This table reports regressions of pool return components on lagged flows. *Flow* is calculated as the daily net of mint and burn values divided by the pool value, expressed as a percentage. The return components are expressed in basis points. See previous Table 7 for variable descriptions. Pool-date observations with absolute value of *Flows* above the 99th percentile are removed.

	(1) Total Return	(2) Fee Yield	(3) ASC	(4) IHR	(5) StakingReturn
SushiAttack	-116.63**	1.95**	-3.18*	-114.83**	-0.00
	(-2.24)	(2.01)	(-1.67)	(-2.25)	(-0.19)
UniIncentive	-25.38*	-4.98***	1.80***	-17.60	-0.07***
	(-1.93)	(-7.69)	(5.59)	(-1.33)	(-15.87)
$Flow_{t-1}$	-1.26**	-0.06**	0.14	-1.47**	0.00
	(-2.16)	(-2.17)	(0.93)	(-2.18)	(1.24)
$Flow_{t-2}$	0.53	-0.02	-0.08*	0.57	-0.00
	(1.21)	(-0.76)	(-1.90)	(1.27)	(-1.25)
$Flow_{t-3}$	0.00	0.01	0.14	-0.17	0.00**
	(0.01)	(0.31)	(1.48)	(-0.34)	(1.97)
$TotalReturn_{t-1}$	0.00				
	(0.04)				
$TotalReturn_{t-2}$	-0.01				
	(-1.26)				
$TotalReturn_{t-3}$	0.01				
	(0.95)				
$FeeYield_{t-1}$		0.30***			
		(3.45)			
$FeeYield_{t-2}$		0.09			
		(1.62)			
$FeeYield_{t-3}$		0.08**			
		(2.12)			
ASC_{t-1}			0.04**		
			(2.19)		
ASC_{t-2}			0.01		
			(0.99)		
ASC_{t-3}			0.00		
			(0.54)		
IHR_{t-1}				0.01	
				(0.44)	
IHR_{t-2}				-0.01	
				(-1.02)	
IHR_{t-3}				0.01	
				(1.05)	
$StakingReturn_{t-1}$					0.02*
					(1.73)
$StakingReturn_{t-2}$					0.01
					(0.77)
$StakingReturn_{t-3}$					0.00
					(0.26)
Constant	1.74	4.80***	-15.71***	9.13***	-0.07***
	(0.67)	(9.01)	(-17.05)	(3.10)	(-9.56)
Obs.	50,019	50,019	50,019	50,019	50,019
R^2	0.000	0.148	0.002	0.000	0.001
Number of pairid	200	200	200	200	200
Pool FE:	Yes	Yes	Yes	Yes	Yes
Date FE:	Yes	Yes	Yes	Yes	Yes

SE clustered by pool and date
T-statistics in parentheses. *p<.05; **p<.01; ***p<.001

7 Market Resiliency Benefits

Algorithmic liquidity providers can (and do) withdraw from markets in turbulent times, (Anand and Venkataraman, 2016). This can exacerbate flash crashes (Kirilenko et al., 2017), contribute to instability in the form of extreme price movements in a number of securities concurrently (Brogaard et al., 2016), and increase the correlation of liquidity across securities (Malceniece et al., 2019).[25] Consequently, the premium placed on liquidity risk has increased in recent years, as found by Pastor and Stambaugh (2019).

AMMs (such as Uniswap) may improve the resiliency of markets through less volatile prices and liquidity. Price volatility may be improved because prices are set in batches, in line with the Ethereum each blockchain update which ranges from 11 to 19 seconds. This batching may prevent 'feedback loops' that occur in crashes and flash crashes on centralized exchanges. Liquidity resiliency may also be improved by the design of AMMs which mutualize losses and gains from liquidity provision amongst all liquidity providers. On centralized exchanges liquidity providers, by contrast, are competitive.

Additionally, on centralized exchanges LPs are incentivized to cancel their resting stale orders to minimize adverse selection on a frequent basis and incur no explicit costs for cancellations — only their queue positions. By contrast, in AMMs LPs incur 'gas fees' both to cancel their liquidity and resubmit it.[26]

I provide evidence that Uniswap AMM prices are far less volatile than centralized markets. I first identify Extreme Price Movements (EPMs) in the sample following the methodology of Brogaard et al. (2016), focusing on EPMs that are transitory — often called 'flash crashes'. I classify 10 second return periods that exceed the 99.9% distribution cutoff as EPMs and those that revert over 75% of the price change within 30 minutes as 'flash crashes'. I focus only on the largest Uniswap pool, WETH-USDT. I perform an event study on the one of the largest of these crashes, plotting prices on Uniswap and Binance during the episode of extreme volatility in Figure 13. The first panel shows a crash on the 2nd of August, 2020 where the price of ETH-USDT falls by 20.5% before recovering 9%.

[25] There are also concerns about deteriorating liquidity in smaller securities, (SEC, 2019), implying that perhaps that the benefits of automation may be limited to large, and thus already relatively liquid securities.

[26] A small competing AMM to Uniswap called "DODO" even levies a penalty for removing liquidity, see: DODO (2020).

By contrast, the fall in Uniswap's price is far more orderly, falling only 9.5% in a manner that excludes rapid falls and rebounds.

The second panel shows measures of liquidity on Binance and Uniswap during the crash. Because cryptocurrency markets have exceedingly small minimum tick sizes, if depth is measured as a fixed number of orderbook levels (25, 10 or 1-level best bid or offer, for example) exhibits extreme volatility, (Dyhrberg et al., 2019). I follow the method proposed by van Kervel (2015) and calculate the sum of Binance orderbook depth levels at each update that are at prices within a threshold of 10 basis points of the midpoint.[27] Uniswap liquidity is calculated as the square root of the pool constant-K, which reflects inflows into the pool from fees and mints as well as outflows from burns.[28] The chart shows that liquidity declines from an average of $300k within 10 basis points before the crash to a range of zero to $20k afterwards. By comparison, Uniswap liquidity increases due to the fee revenues from trades that occur during the crash, with no withdrawals of liquidity. Uniswap shows a marked improvement in both the volatility and resiliency of liquidity between the two exchanges during the crash.

[27] This threshold is set by taking the 99.9th percentile of the 25th depth price.

[28] Alternative approaches to plotting the pool-value would be significantly impacted by changes in the pool price.

Figure 13: Uniswap Price vs Binance Midpoint During Flash Crashes
This figure plots prices on Uniswap and Binance for one of the largest 'flash crashes' in ETH-USDT over the sample period, the 2nd of August, 2020 in the top Panel and liquidity measures during the same period in the bottom panel. The midpoint is used as the price on Binance and the WETH-USDT pool price is used as the price on Uniswap. Binance liquidity is measured as the total value of orderbook depth within 10 basis points of the midpoint, expressed in thousands USD. The Uniswap liquidity is measured as the square root of the pool constant-K, expressed in thousands.

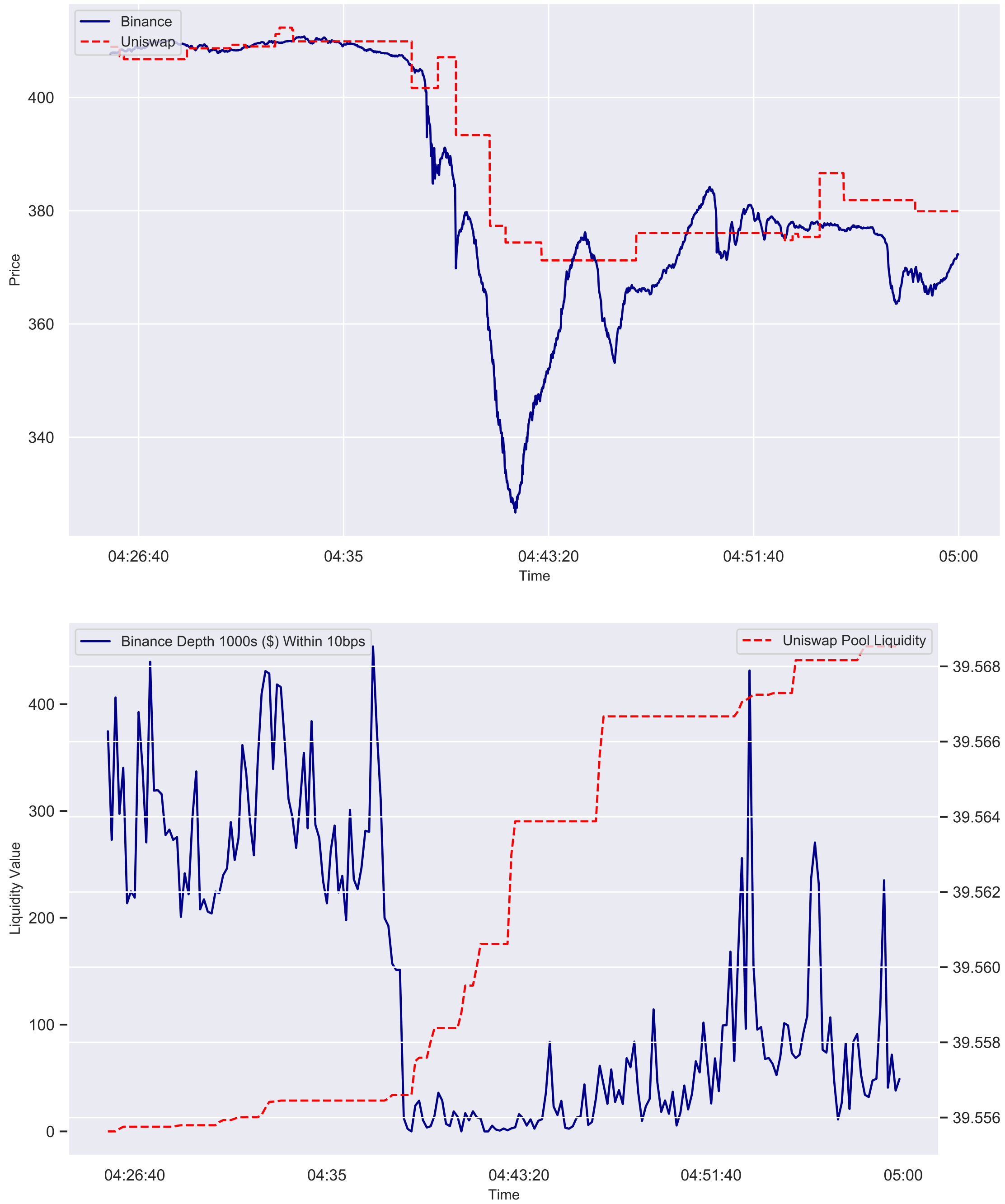


Table 9 sets out return and liquidity statistics for Uniswap and Binance during the 231 non-overlapping flash crashes identified in the sample, covering 1-minute before the crash and 30-minutes afterwards. The results are consistent with the event study above. The table shows that Uniswap experiences less price and liquidity volatility than Binance. Absolute returns are 35% lower on average and the variance in prices is 12% lower on average. The change in liquidity is measured as the change in 25-level depth each minute for Binance and the net of mints and burns as a percentage of the pool size in each minute for Uniswap. Changes in Uniswap liquidity are extremely rare, the mean change is less than a basis point each minute (or around three blocks). The liquidity variance shows that the pool only ever increases in size during flash crashes, typically benefiting from fee revenue significantly — mean pool size increases by 0.35% and 6.7% at the 99th percentile. Therefore, the Uniswap AMM experiences essentially no liquidity withdrawal during extreme periods of volatility across the entire sample. This provides evidence that liquidity is significantly more resilient in AMMs and prices less volatile during periods of market stress.

Table 9: Return and Liquidity Metrics within Flash Crashes on Binance vs Uniswap
This table presents return and liquidity metrics during each 30-minute flash crash period for Binance and Uniswap separately. A flash crash period is identified as a 99.9th percentile negative return on Binance that also reverts more than 75% in the subsequent 30-minute period. *Returns* are calculated each minute as $|R_t|$ and R_t, *Price Variance* as $(m_{max} - m_{min})/m_0$ where the m_0 is the midpoint 1-minute before the crash. For Uniswap, *Returns* are calculated using the pool price. Δ *Liquidity* is calculated as the change in depth or the % net mint or burn in the Uniswap pool each minute as a proportion of the pool size $|\Delta d|$ and Δd. *Liquidity Variance* is calculated as difference between the max and min 25-level depth during the the 30-minute period divided by the depth before the crash, $(d_{max} - d_{min})/d_0$. For Uniswap, the pool-constant K is used. All measures are in percent.

	Uniswap					**Binance**				
	Mean	std	p1	p50	p99	Mean	std	p1	p50	p99
Absolute Returns	0.209	0.309	0.000	0.083	1.337	0.321	0.384	0.003	0.212	1.743
Returns	-0.011	0.373	-1.180	0.000	1.052	-0.007	0.500	-1.427	0.000	1.302
Price Variance	2.134	1.128	0.000	1.856	5.525	2.433	1.181	0.681	2.153	5.525
Absolute Δ Liquidity	0.005	0.158	0.000	0.000	0.002	38.985	111.799	0.336	21.832	328.336
Δ Liquidity	0.002	0.159	0.000	0.000	0.000	14.965	117.452	-77.540	0.000	328.336
Liquidity Variance	0.349	1.642	0.000	0.013	6.720	266.213	424.419	52.408	154.929	2326.003

8 Conclusion

Dealers or market makers have performed the valuable function of liquidity provision in financial markets for centuries. Remarkably, it is possible to replace them with a simple set of code. In this paper I show that the architecture created by this code, an Automated Market Maker, is able to feasibly perform the two main functions of financial markets — it is able to efficiently provide liquidity while also setting accurate prices. It does this by charging fees to compensate 'liquidity providers' that provide capital (or staked assets) to this piece of code. While these fees are fixed for those trading with the AMM, capital flows to discrete AMM "pools" equilibrate the return on capital to appropriately compensate the liquidity providers for their adverse selection risks. I characterize and derive the components of these mechanisms and show evidence of this equilibrium to evidence the viability of this new market design. I show that these AMMs ultimately produce prices that are remarkably similar to their centralized counterparts, and are thus represent viable alternatives.

My evidence suggests that pools have difficulty reaching equilibrium, and remain unprofitable in high volatility pairs (which are often smaller pools) as well as in most token-token pairs. The resolution of these negative fees is that some liquidity providers endogenously depart, reducing the sizes of these pools and consequently increasing the proportional fee revenue. Such an equilibrating mechanism necessarily reduces the capacity of the AMM to provide large levels of liquidity. In traditional markets, by contrast, market-makers can vary both the *price* of liquidity (by charging a higher spread), as well as varying the amount of liquidity (by varying the depth of their limit orders). In this way, the *amount* of liquidity, or the market depth, can remain constant. This suggests that a dynamic (as opposed to a fixed) fee for liquidity providers would improve overall liquidity in the AMM. Indeed, the latest iterations of AMMs are experimenting with the availability of multiple pools for each pair of assets, where each of the pools carry different fixed fees for liquidity provision. This theoretically provides LPs with an additional mechanism for equilibration — the ability to vary the 'price' of liquidity without varying the quantity.

It is important to emphasize, however, that this paper examines a project that is only a few years old, that has and will continue to undergo further improvements over time. This paper examines only the second iteration of Uniswap's design. It is remarkable that even at such an early stage AMMs offer tangible benefits in comparison to centralized markets, which have had a few centuries head start. There are still several limitations, but rapid iterations to the design of the AMM (absent typical regulatory barriers) may quickly close this gap. While modifying the AMM is as simple as modifying the code, the economic incentives and equilibrating mechanisms implicit in such changes require careful consideration.

Appendices

A Framework of AMM Properties and LP Returns

Consider an AMM for the asset pair USD and ETH, where x_t and y_t are the quantities of USD and ETH in the AMM at time t, respectively.[29]

Liquidity providers (LPs) stake a total of x_0 of USD and y_0 of ETH at $t = 0$ when the price of ETH in dollars is P_0. This gives the pool the constant $K = x_0 y_0$. The total USD value of the assets in the AMM at $t = 0$ is V_0, which (as I show later) is equal to $2x_0$ or twice the amount of USD deposited in the AMM, because the design of the AMM ensures that the value of both assets in the AMM is equal at the time of staking assets and at any subsequent time.

The AMM facilitates trades by liquidity demanders (LDs or "traders"). Trades with the AMM occur in discrete time at $t = 1, 2, 3, \ldots T$ where $t = T$ is the time at which the LP ceases liquidity provision and exits the game with the assets that can be withdrawn from the AMM at the time.

A trade is an asset swap in which one asset is added and the other removed in quantities such that the constant K is maintained (by design of constant product AMMs). Thus, a trade at time t that buys ETH involves removing $\Delta y_t = y_t - y_{t-1}$ from the AMM and adding in $\Delta x_t = x_t - x_{t-1}$ to the AMM such that the constant is preserved, $x_t y_t = K = x_{t-1} y_{t-1} = (x_{t-1} - \Delta x_t)(y_{t-1} + \Delta y_t)$. Conversely, a trade at time t that sells ETH by adding Δy_t to the AMM and removing Δx_t from the AMM, $x_t y_t = (x_{t-1} - \Delta x_t)(y_{t-1} + \Delta y_t)$.

Each trade also pays a fee equal to 30 basis points (bps) of the trade value, $f_t = 0.003 x_t$. While fees would usually be added to the liquidity pool thereby slightly increasing the pool's K with each trade, for expositional clarity and tractability, I assume that fees accrue to a separate account, which is equivalent to assuming that LPs withdraw the fees from the AMM to maintain their staked amount of liquidity constant and rather than reinvest the fees to increase the provided liquidity. Over the time horizons in question,

[29] I use USD and ETH throughout an example of two assets staked in the AMM because this is the most popular combination of assets in AMMs. The derivations and mechanics of liquidity provision are the same for any pair. One asset is the unit of account (or numeraire), in this case USD is considered the unit of account, and the other is denoted as the "risky asset" as its price can fluctuate relative to the unit of account. While it is natural to think of USD as "stable" and ETH being a risky asset that fluctuates in price, ETH can also be considered the unit of account without affecting the derivations in this paper.

this simplifying assumption makes little difference to the overall results. If (as I will show later) there is an equilibrium level of liquidity (value of staked assets), then maintaining that equilibrium would require an LP to withdraw fees, consistent with the approach I use in the derivations below.

While the amount of liquidity staked in an AMM can be varied through time by LPs "minting" (staking more of both assets) and "burning" (redeeming or withdrawing assets), I assume that T is a sufficiently short period such that no mints or burns occur during the period from $t = 0$ to $t = T$. Allowing for mints or burns during this staking horizon would not qualitatively change the key results, it merely complicates the calculations.

Defining some trading volumes that will be used in later calculations, during the staking period $t = 0$ to $t = T$, the total dollar volume of LD buys is $BUYS_T = \sum_{t=1}^{t=T} max[0, \Delta x_t]$, i.e., the total amount of USD added to the AMM, and similarly the total dollar volume of LD sells is $SELLS_T = \sum_{t=1}^{t=T} max[0, -\Delta x_t]$, that is, the total amount of USD removed from the AMM. The total dollar volume of trading is $Q_t = BUYS_T + SELLS_T = \sum_{t=1}^{t=T} |\Delta x_t|$. The balanced (roundtrip) order flow is $BAL_T = 2 * min[BUYS_T, SELLS_T]$ and the order imbalance between buys and sells is $OIB_T = BUYS_T - SELLS_T = \sum_{t=1}^{t=T} \Delta x_t$, such that the total dollar volume is the sum of the balanced volume and the absolute value of the order imbalance $Q_T = BAL_T + |OIB_T|$. It is convenient to also normalize the trade volumes by the total pool value, V_0, such that the turnover in horizon T is $TURN_T = Q_T/V_0$, the balanced turnover ("relative" balanced volume) is $RBAL_T = BAL_T/V_0$, and the relative order imbalance is $ROIB_T = OIB_T/V_0$.

B AMM Properties

PROPERTY 1: Ignoring fees, the trade prices in this AMM ($ per ETH) for any quantity of ETH Δy_t are given as follows:

- Consider a trade to buy $\Delta y_t > 0$ units of ETH in exchange for paying $\Delta x_t > 0$ To retain the constant product, I must have $(x_{t-1} + \Delta x_t)(y_{t-1} - \Delta y_t) = K = x_t y_t$. Rearranging, the price of the swap (how many dollars are spent per unit ETH in the swap) is:

$$P(\Delta y_t) = \frac{\Delta x_t}{\Delta y_t} = \frac{x_{t-1}}{y_{t-1} - \Delta y_t} \tag{1}$$

- Similarly, a trade to sell $\Delta y_t'$ units of ETH and receive $\Delta x_t'$ will occur at the price (how many dollars are received by the trader per unit ETH in the swap):

$$P(\Delta y_t) = \frac{\Delta x_t}{\Delta y_t} = \frac{x_{t-1}}{y_{t-1} + \Delta y_t}$$

which is the same as 1, just denoting the sell quantity as a negative value, $\Delta y_t = -\Delta y_t^{'} < 0$, that is, the AMM is governed by the price function (1) in which I have $\Delta y_t > 0$ for buys and $\Delta y_t < 0$ for sells of ETH.

- This price function gives the AMM's bonding curve in Figure 1.

Property 2: The 'midpoint' ETH price of the pool in USD (the price of an infinitesimally small trade that has negligible price impact) is purely a function of the two asset quantities in the pool at the time: $P_{0,MID} = x_0/y_0$ or more generally $P_{t,MID} = x_t/y_t$.

- To see this, consider an infinitesimally small swap to buy Δy_t units of ETH and pay Δx_t in the pricing function (1) above, that is, as $\Delta y_t \to 0, P_(\Delta y_t) \to \frac{x_t}{y_t}$.

Property 3: Ignoring fees, the sequence in which trades occur does not matter for the final outcome (state) of the AMM, being its pool quantities and midpoint price.

- To see this, consider a trade to buy or sell $\Delta y_t^{'}$ units of ETH ($\Delta y_t > 0$ implies a buy, and $\Delta y_t < 0$ implies a sell), followed by a trade to buy or sell $\Delta y_t''$ units of ETH ('first scenario') and compare that with the two trades occurring in the reverse sequence ($\Delta y_t''$ first and $\Delta y_t'$ second, 'second scenario'). Let the \$ paid or received in each of the trades be $\Delta x_t^{'}$ and $\Delta x_t''$ in Scenario 1 and Δy_t^{**} and Δy_t^{*} in scenario 2 (when buying ETH $\Delta x > 0$ is the amount of \$ paid by the trader, and when selling ETH $\Delta x < 0$ is the amount received).
- To retain the constant product, in Scenario 1 I must have: $(x_0 + \Delta x')(y_0 - \Delta y') = K$, giving pool quantities $x_1 = x_0 + \Delta x'$ and $y_1 = y_0 - \Delta y'$ after 1^{st} trade

  $(x_1 + \Delta x'')(y_1 - \Delta y'') = K$, giving pool quantities $x_2 = x_1 + \Delta x''$ and $y_2 = y_1 - \Delta y''$ after 2^{nd} trade

  and thus, $x_2 = x_0 + \Delta x' + \Delta x''$ and $y_2 = y_0 + \Delta y' + \Delta y''$ and

  $(x_0 + \Delta x' + \Delta x'')(y_0 - \Delta y' - \Delta y'') = K$
- Similarly, in Scenario 2 (reverse order) I must have:

  $(x_0 + \Delta x^{**})(y_0 - \Delta y'') = K$, giving pool quantities $x_1 = x_0 + \Delta x^{**}$ and $y_1 = y_0 - \Delta y''$ after 1^{st} trade

$(x_1 + \Delta x^*)(y_1 - \Delta y') = K$, giving pool quantities $x_2 = x_1 + \Delta x^*$ and $y_2 = y_1 - \Delta y'$ after 2^{nd} trade

and thus, $x_2 = x_0 + \Delta x^* + \Delta x^{**}$ and $y_2 = y_0 + \Delta y' + \Delta y''$ and

$(x_0 + \Delta x^* + \Delta x^{**})(y_0 - \Delta y' - \Delta y'') = K$

- Equating the two final equations in each scenario, I get:

$$\left(x_0 + \Delta x' + \Delta x''\right)\left(y_0 - \Delta y' - \Delta y''\right) = \left(x_0 + \Delta x^* + \Delta x^{**}\right)\left(y_0 - \Delta y' - \Delta y''\right)$$

$\Delta x' + \Delta x'' = \Delta x^* + \Delta x^{**}$ implying x_2 is the same under both scenarios and so is y_2

- Therefore, under both trade sequences, the final quantities of the two assets in the AMM are the same, and so too must be the final 'midpoint' prices of the AMM, so the sequence in which trades occur does not matter for the final outcome (state) of the AMM.

Property 4: Still ignoring fees, a roundtrip trade reverts the price back to the original, reverts the pool quantities back to the original, and the trader breaks even (receives the same amount of \$ as she paid).

- To see this, consider a trade to buy Δy units of ETH in exchange for paying Δx (pay $\$\Delta x$) and then selling the same Δy units of ETH back to receive $\Delta x'$. To retain the constant product, I must have

$$\left(x_0 + \Delta x - \Delta x'\right)(y_0 - \Delta y + \Delta y) = K = x_0 y_0$$

$$\left(x_0 + \Delta x - \Delta x'\right)(y_0) = x_0 y_0$$

$$x_0 + \Delta x - \Delta x' = x_0$$

$$\Delta x - \Delta x' = 0$$

$$\Delta x = \Delta x'$$

- Therefore, the dollars paid to the pool equal the dollars received from the pool, so the pool quantities all revert back to their original, so too must the price.

Property 5: Two small trades in the same direction are equivalent to one larger trade in the same direction with quantity equal to the sum of the two smaller quantities (same end price in the AMM, same end state in terms of quantities in the AMM, and same cost to the trader). In other words, there is no advantage from trade slicing as the outcomes are the same.

- To see this, consider a trade to buy Δy units of ETH, then buy another $\Delta y'$ units of ETH (first scenario) and compare that with a trade to buy $\Delta y'' = \Delta y + \Delta y'$ units of ETH (second scenario). Let the \$ paid in each of the trades be $\Delta x, \Delta x'$, and $\Delta x''$.
- To retain the constant product, I must have

$$\left(x_0 + \Delta x + \Delta x'\right)\left(y_0 - \Delta y - \Delta y'\right) = K \text{ (Scenario 1)}$$
$$\left(x_0 + \Delta x''\right)\left(y_0 - \Delta y''\right) = K \text{ (Scenario 2)}$$

Thus the two left hand sides must be equal,

$$\left(x_0 + \Delta x + \Delta x'\right)\left(y_0 - \Delta y - \Delta y'\right) = \left(x_0 + \Delta x''\right)\left(y_0 - \Delta y''\right)$$

and recognizing that $\Delta y'' = \Delta y + \Delta y'$, I get:

$$\Delta x'' = \Delta x + \Delta x'$$

which says that the trader would pay just as much in \$ for the one big purchase of ETH vs the two small buys that sum to the big trade's volume. Therefore, also the quantities of both assets left in the pool are the same under the two scenarios, and thus so too is the ending "midpoint" price in the pool (even though the trade prices are different).

Property 6: If the AMM receives a series of trades, only the buy/sell imbalance quantity of the series of trades is needed to work out the impact on the AMM's state (change in asset quantities in the pool and the pool's midpoint price), that is, the balanced part of volume (buy volumes equal to sell volumes) has no impact (irrespective of what combination of trades makes the balanced volume), the sequence of trades does not matter, and the AMM is "memoryless" in that the impact of a trade depends only on the current state of the AMM and not the history of trades.

- To see this, exploit Property 3 saying the sequence does not matter and Property 5 saying I can sum trade quantities together, and sum all the ETH buy quantities to the aggregate quantity $\Delta y_{BUY} = \sum_{i \in BUYS} [\Delta y_i] \geq 0$ and sum all the ETH sell quantities to the aggregate quantity $\Delta y_{SELL} = \sum_{i \in SELLS} [-\Delta y_i] \geq 0$ (note here I am defining the sell quantity as a positive value, but in the pricing function (1) I would have $\Delta y = -\Delta y_{SELL}$). Let the aggregate \$ spent by the trader on the buys be $\Delta x_{BUY} \geq 0$ and the aggregate \$ received by the trader on the sells be $\Delta x_{SELL} \geq 0$. The total trade volumes can be broken up into the order imbalance, $\Delta y_{OIB} = \Delta y_{BUY} - \Delta y_{SELL}$ and the roundtrip trades, $\Delta y_{ROUNDTRIP} = 2 \times \min\left[\Delta y_{BUY}, \Delta y_{SELL}\right]$. For example, if I have buys that sum to 2ETH and sells that sum to 5ETH then the 2 ETH buys are perfectly offset by 2ETH in sells (total roundtrip volume of 4ETH) but the additional 3ETH in sells is not offset so the order imbalance is -3.

- From Property 4 I know that the roundtrip trades have no effect on the state of the AMM, so only the imbalance matters. Thus, any sequence of any number of trades $t = 1, \ldots N$ in any directions, can be summarized in a sufficient statistic, $\Delta y_{OIB} = \Delta y_{BUY} - \Delta y_{SELL}$, which determines the change in the AMM asset quantities and midpoint price from x_0, y_0, P_0 to: $y_N = y_0 - \Delta y_{OIB}, x_N = x_0 + \frac{x_0 \Delta y_{OIB}}{y_0 - \Delta y_{OIB}}$ (drawing on equation (1)), and

$$\begin{aligned}P_{N,MID} = \frac{x_N}{y_N} = \frac{x_0 + \frac{x_0 \Delta y_{OIB}}{y_0 - \Delta y_{OIB}}}{y_0 - \Delta y_{OIB}} &= \frac{x_0}{y_0 - \Delta y_{OIB}} + \frac{x_0 \Delta y_{OIB}}{(y_0 - \Delta y_{OIB})^2} \\ &= \frac{x_0}{y_0 - \Delta y_{OIB}} \left(1 + \frac{\Delta y_{OIB}}{y_0 - \Delta y_{OIB}}\right)\end{aligned}$$

- The midpoint price change in the AMM, expressed as a return, is:

$$\begin{aligned}R - 1 = \frac{P_{N,MID}}{P_{0,MID}} - 1 = \left(\frac{y_0}{x_0}\right)\left(\frac{x_0}{y_0 - \Delta y_{OIB}}\right)\left(1 + \frac{\Delta y_{OIB}}{y_0 - \Delta y_{OIB}}\right) \\ = \frac{y_0}{y_0 - \Delta y_{OIB}}\left(1 + \frac{\Delta y_{OIB}}{y_0 - \Delta y_{OIB}}\right) - 1 \quad (9)\end{aligned}$$

  that is, in any series of trades, the change in the price of ETH in the AMM is purely a function of the order imbalance quantity (how much more ETH was bought than sold) and the initial ETH quantity in the pool, which determines the pool depth.

PROPERTY 7: At every point in time, including when assets are staked and redeemed, assuming the 'midpoint' ETH price of the pool is approximately equal to the value of ETH (arbitrage has driven the AMM to the 'correct' price) the value of each of the two assets staked by the LP are equal, when measured in one unit of account, for example, the USD value of x_t is always equal to the USD value of y_t

- To see this, convert the quantity of asset y into an equivalent value in currency x using the midpoint price implied by the AMM at the time: $y_t P_{t,MID} = y_t \frac{x_t}{y_t} = x_t$
- It follows that an LP stakes assets in equal value weights, and those equal value weights are maintained by the AMM pricing function even as trades occur.

C Effects of fees

Fees, if left in the AMM (default option), result in an increase in the pool's assets by the amount of the fee each time a transaction occurs, effectively slowly increasing the constant K through time due to reinvestment by the LPs. In such cases, some of the properties above will not hold exactly, but will be a reasonable approximation. For example, consider a buy followed by an equal size sell. The second trade, the sell, will have a slightly smaller price impact due to the slight increase in K from the first trade. Thus, the AMM midpoint price will not return exactly to the same price following the roundtrip trade (but it will return to a very close value), so Properties 3-6 only hold to an approximation. In practice, the approximation will be very close, irrespective of whether fees are left in the pool as reinvested liquidity or withdrawn, given fees are 0.30% of a transaction value and typical staking horizons are only a matter of a few days ¡example . . . ¿. Treating the fees as if they accrue to a separate account (equivalent to withdrawing the fees) maintains tractable solutions to the various derivations below. I therefore adopt this approach throughout the rest of the document, treating fees as if they accrue to an account but are not 'recycled' back into the liquidity pools.

D Digging in to the LP return components

I now examine mathematically what drives the three return components of an LP's returns. I start by characterizing how the return components relate to price changes of the staked assets and the volatility of the staked asset returns. After that, I express the return components as functions of the order flow to the AMM and the AMM pool size. Starting with inventory holding returns, first recognize that this is the return to a portfolio of two assets (USD and ETH in our carried through example) where the quantities of those two assets do not change through time and the return to a portfolio of assets is the value-weighted average of the individual asset returns. Second, recognize that the weights on the two assets are $\frac{1}{2}$ and $\frac{1}{2}$ due to the mechanics of the constant product function in the AMM (see Property 7 above). Third, recognize that the return to a constant quantity of an asset as measured in units of that asset is zero (for example, return on the USD staked in the pool is zero because the USD price of 1 USD is constant at 1). Therefore, I have:

$$IHR_T = \frac{V_{T,FIXED}}{V_0} - 1 \tag{10}$$

$$IHR_T = \frac{1}{2}\left(\frac{P_T}{P_0} - 1\right) = \frac{1}{2}r_y \tag{4}$$

where $r_y = \frac{P_T}{P_0} - 1$ is the return on asset y (e.g. ETH). Thus, the inventory holding return is just a dampened return on the risky asset.

It follows that inventory holding risk, the expected volatility of the inventory holding returns, is one quarter of the risky asset's volatility (σ_y) :

$$\sigma_{IHR} = \mathrm{Var}\left(\frac{1}{2}r_y\right) = \frac{1}{4}\sigma_y \tag{4C}$$

The inventory holding return can also be expressed as a function of order flow, given price changes are closely linked to order flow by design in the AMM. To do this, I invert the pricing function of the AMM so that instead of getting how a price will change given a trading volume, I instead infer what must be the trading volume for a given price change. Equation (11) below [4] arrives at the following:

$$x_T = x_0\sqrt{P_T/P_0} \quad \text{and} \quad y_T = y_0\sqrt{P_0/P_T} \tag{11}$$

Taking the first of these equations and recognizing that $x_T = x_0 + \Delta x_T$ where $\Delta x_T = \sum_{t=1}^{t=T} \Delta x_t = OIB_T$ is the USD value of the order imbalance during the staking period (the sum of all the USD received by the AMM from traders buying ETH minus all the USD paid out by the AMM to traders selling ETH during the period (trades) $t = 1$ to $t = T$), I have:

$$\frac{P_T}{P_0} = \left(\frac{\Delta x_T}{x_0} + 1\right)^2 \tag{11B}$$

which says that the price of ETH is increasing ($P_T > P_0$) with positive signed order flow for ETH because $\Delta x_T > 0$ implies there has been a greater buy volume for ETH from traders than sell volume and thus the AMM has increased its holding of USD. Substituting (11B) into (4) and replacing[30] with $\frac{V_0}{2}$ I get:

$$
\begin{aligned}
IHR_T &= \frac{1}{2}\left[\left(2\frac{\Delta x_T}{V_0}+1\right)^2-1\right] &\text{(4D)}\\
&= \frac{1}{2}\left[(2ROIB_T+1)^2-1\right] &\text{(4E)}\\
&= 2ROIB_T + 2ROIB_T^2 &\text{(4F)}
\end{aligned}
$$

What equation 4F shows is that the inventory holding return is an increasing function of the order imbalance (buy volume minus sell volume) for ETH, expressed in USD terms, as a proportion of the total pool value (USD), that is, $\frac{\Delta x_T}{V_0} = ROIB_T$, or "relative order imbalance".[31] The intuition is that an LP is ultimately long the risky asset, so buying pressure for the risky asset is ultimately going to contribute a positive return to the LP's total return from the increase in the price of the risky asset.

Second, considering the adverse selection costs, I have:

$$
\begin{aligned}
ASC_T &= \left(\frac{V_T}{V_0}-1\right)-\left(\frac{V_{T,FIXED}}{V_0}-1\right) &(12)\\
&= \frac{V_T - V_{T,FIXED}}{V_0} &(13)\\
&= \frac{x_T + y_T P_T - (x_0 + y_0 P_T)}{x_0 + y_0 P_0} &(14)\\
&= \frac{(x_T - x_0) + (y_T - y_0)\, P_T}{v_0} &(15)
\end{aligned}
$$

I pause here to note some nice intuition from the above expression, which illustrates how the adverse selection cost is driven by changes in the asset quantities within the AMM.

[30]Because,

$$V_0 = x_0 + y_0 P_0 = x_0 + y_0 \frac{x_0}{y_0} = 2x_0 \Rightarrow x_0 = \frac{V_0}{2}$$

[31]To be clear, $ROIB_T$ is the sum of all the USD received by the AMM from traders buying ETH minus all the USD paid out by the AMM to traders selling ETH during the staking period (trades) $t = 1$ to $t = T$, divided by the initial (at $t = 0$) USD value of the staked assets.

The adverse selection cost is the change in the quantity of USD left in the pool, $x_T - x_0$, plus the change in the quantity of ETH left in the pool, expressed in USD by multiplying by the final price of ETH, $(y_T - y_0)\, P_T$. If traders using the AMM during the staking period from $t = 1$ to $t = T$ were net buyers of ETH from the AMM, the USD quantity will have increased, $(x_T - x_0) > 0$, and the ETH quantity will have decreased $(y_T - y_0) < 0$ and, importantly, due to the buying the ETH price will have increased $P_T > P_0$ which means the LP ends up with a smaller quantity of the asset that has appreciated in value. Conversely, if traders using the AMM during the staking period were net nave decreased $P_T < P_0$ which means the LP ends up with a lager quantity of the asset that has fallen in value. This tendency for the LP to be left with more of the risky asset when it is worth less and less of the risky asset when it is worth more is precisely what gives rise to the adverse selection cost. Continuing the derivation, (15) can be re-expressed[32] as:

$$ASC_T = \frac{(x_T - x_0) + \left(\frac{x_T}{P_T} - \frac{x_0}{P_0}\right) P_T}{2x_0} \tag{16}$$

$$= \frac{2x_T - x_0 \left(1 + \frac{P_T}{P_0}\right)}{2x_0} \tag{17}$$

Now recognize that because the prices in the AMM are a function of the asset quantities in the AMM, I can invert the pricing equation to express changes in the quantities of assets in the pool as a function of changes in the pool prices. Thus, I can rearrange the constant product equation that governs the AMM's pricing function, $x_T y_T = K = x_0 y_0$, to express how x_T and y_T evolve from x_0 and y_0 as the AMM's midpoint prices change:

$$x_T = x_0 \sqrt{P_T/P_0} \quad \text{and} \quad y_T = y_0 \sqrt{P_0/P_T} \tag{18}$$

Substituting the first equation in (18) into (17) and rearranging, I get the adverse selection cost expressed in terms of price changes:

[32]Recognizing that midpoint prices are given by $P_t = \frac{x_t}{y_t} \Rightarrow y_t = \frac{x_t}{P_t}$ and $V_0 = x_0 + y_0 P_0 = 2x_0$ and substituting these into the equation.

$$ASC_T = \sqrt{\frac{P_T}{P_0}} - \frac{1}{2}\left(\frac{P_T}{P_0} + 1\right) \tag{19}$$

$$= \sqrt{R_T} - \frac{1}{2}\left(R_T + 1\right) \leq 0 \tag{5}$$

where $R_T = \frac{P_T}{P_0}$ is the ratio of prices of the risky asset at the end of the staking period $t = T$ compared to the beginning, $t = 0$. Equation (5) shows that the adverse selection cost is a function of how prices change from their initial values. Taking the derivative, $\frac{d(ASC_T)}{dR_T} = \frac{1}{2\sqrt{R_T}} - \frac{1}{2}$, I see that ASC_T has a maximum value of 0 when $R_T = 1$, i.e.,

(i) The adverse selection cost component of an AMM LP's returns are always less than or equal to zero, that is, a true "cost";

(ii) Are minimized when asset prices do not change $(R_T = 1 \Rightarrow P_T = P_0)$; and

(iii) Increase in both directions of asset price changes, i.e., as the price of the risky asset increases or decreases from its initial value, the LP loses money on adverse selection.

I can also express the adverse selection cost as a function of order flow, given price changes are closely linked to order flow by design in the AMM. To do this, I use the previously inverted pricing function of the AMM (equation 11B) and substitute it in for the price ratio $\left(\frac{P_T}{P_0}\right)$ in (5) to get:

$$ASC_T = \frac{\Delta x_T}{x_0} + 1 - \frac{1}{2}\left(\left(\frac{\Delta x_T}{x_0} + 1\right)^2 + 1\right) \tag{5C}$$

$$= -2\left(\frac{\Delta x_T}{V_0}\right)^2 = -2\left(ROIB_T\right)^2 \leq 0 \tag{5D}$$

where $ROIB_T = \frac{\Delta x_T}{V_0}$ is the "relative order imbalance" (ETH buy volume minus ETH sell volume, both expressed in USD, as a proportion of the total USD pool value at the start of the staking period). Therefore, equation (5C) shows that the adverse selection cost is a negative quadratic function of the order imbalance, i.e.,

(i) Balanced or roundtrip trades do not contribute to the ASC_T, only the imbalance between buys and sells;

(ii) ASC_T takes its maximum value of zero when $ROIB_T = 0$, i.e., when there is no imbalance between buys and sells; and

(iii) The ASC_T becomes a larger magnitude cost as the order imbalance increases in either direction (more buys than sells or more sells than buys) at a quadratic rate.

Third, the fee yield is unrelated to price changes and is instead a function of the total trade volume, Q_T, irrespective of trade direction and whether the flow is balanced or unbalanced:

$$FY_T = \frac{F_T}{V_0} = \frac{0.003 \sum_{t=1}^{t=T} |\Delta x_t|}{V_0} = \frac{0.003 Q_T}{V_0} = 0.003 TURN_T \tag{20}$$

Thus, higher volumes of trade are expected to increase fee yield. Also, lower pool value (lower V_0) will increase the fee yield. So, ultimately, fee yield is an increasing function of trading volume normalized by pool size, known as pool turnover, $TURN_T = Q_T/V_0$.

I now re-combine the three return components to see the drivers of total returns:

$$R_{TOTAL} = \underbrace{\left(\frac{V_{T,FIXED}}{V_0} - 1\right)}_{\text{Inventory Holding Return}} + \underbrace{\left[\left(\frac{V_T}{V_0} - 1\right) - \left(\frac{V_{T,FIXED}}{V_0} - 1\right)\right]}_{\text{Adverse Selection Cost}} + \underbrace{\frac{F_T}{V_0}}_{\text{Fee Yield}} \tag{21}$$

In terms of price changes:

$$R_{TOTAL} = \underbrace{\frac{1}{2}\left(\frac{P_T}{P_0} - 1\right)}_{\text{Inventory Holding Return}} + \underbrace{\sqrt{\frac{P_T}{P_0}} - \frac{1}{2}\left(\frac{P_T}{P_0} + 1\right)}_{\text{Adverse Selection Cost}} + \underbrace{0.003\frac{Q_T}{V_0}}_{\text{Fee Yield}} \tag{21B}$$

$$R_{TOTAL} = \sqrt{\frac{P_T}{P_0}} - 1 + 0.003\frac{Q_T}{V_0} \tag{21C}$$

In terms of order flow:

$$R_{TOTAL} = \underbrace{2ROIB_T + 2ROIB_T^2}_{\text{Inventory Holding Return}} + \underbrace{-2\left(ROIB_T\right)^2}_{\text{Adverse Selection Cost}} + \underbrace{0.003\frac{Q_T}{V_0}}_{\text{Fee Yield}} \tag{21D}$$

$$R_{TOTAL} = 2ROIB_T + 0.003\frac{Q_T}{V_0} = 2ROIB_T + 0.003 TURN_T \tag{21E}$$

where Q_T to be the total dollar volume of trades during the staking period, $Q_T = \sum_{t=1}^{t=T} |\Delta x_t|$ $TURN_T = \frac{Q_T}{V_0}$ is the pool's turnover in horizon T, and $ROIB_T = \left(\sum_{t=1}^{t=T} \max\left[0, \Delta x_t\right] - \sum_{t=1}^{t=T} \max\left[0, -\Delta x_t\right]\right)/V_0$ is the signed relative order imbalance received by the AMM. Recall that the total dollar volume of trades and turnover can be broken up into the balanced dollar volume and the absolute order imbalance, i.e.,

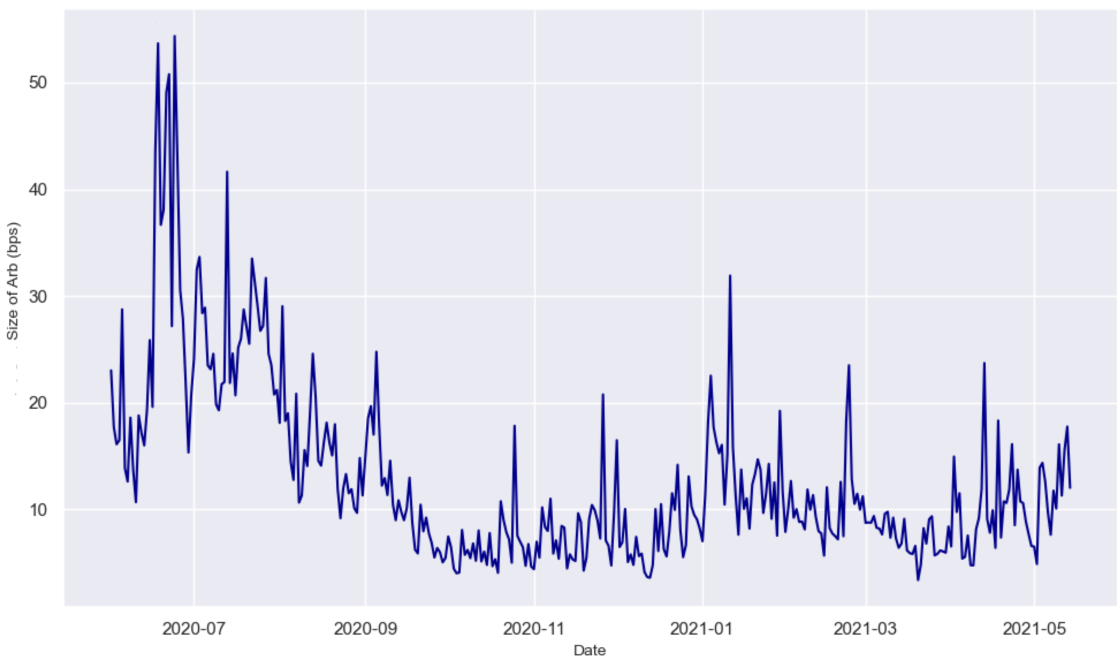
$\frac{Q_T}{v_0} = RBAL_T + |ROIB_T|$. Substituting this decomposition of $\frac{Q_T}{V_0}$ into (4E) I get:

$$R_{TOTAL} = 2ROIB_T + 0.003\,|ROIB_T| + 0.003RBAL_T \tag{21F}$$

Interestingly, while the adverse selection costs are increasing in the absolute amount of order imbalance in either direction (excess buying or excess selling), the total return is directionally increasing in the volume or buys (decreasing in the volume of sells) due to the strong effects of the IHR component (21F). The same effect can be seen in terms of returns: while the ASC increases irrespective of the direction of returns as long as the asset price changes, the total return is positively related to the return to the risky asset (21C).

E Price Efficiency Charts

Figure 14: Size of Actionable and Profitable Arbitrage Opportunities
This figure plots the average size of arbitrage opportunities that remain once trading fee frictions to arbitrage are removed. Arbitrages are calculated between WETH-USDT on Uniswap, the largest pool and Binance ETH-USDT. Arbitrages must persist for more than 19 seconds and must be greater than 40bps worth of frictions.



F Statistics on Liquidity Provider Holding Times

The resting times of LPs is of interest as it informs the appropriate time horizon to assess LP profitability. Resting times are measured as the time interval between the mint and burn events of a given LP. Table 10 shows that LPs provide liquidity for 5.6 days on average, with over 50% of LPs providing liquidity for over 2 days. This proportion is even higher when LPs that never cancel are included in the sample, where I assume the holding time is until the end of the sample. The mean holding time increases to 35.7 days with over 50% holding more than 7.7 days.

Table 10: LP Mint and Burn Statistics and Resting Times in Days
These are calculated by LP for Mint and Burn statistics. For resting times, statistics for each discrete mint and burn event within LPs, across all LPs are calculated. For the top 200 pools by pool value in the sample.

	Mint and Burn Counts	Resting Time (Days) w/ Never Cancels	No Never Cancels
Mean	5.6	35.67	20.32
Median	2.0	7.65	5.29
Std. Dev	55.4	61.57	34.51
p10	1.0	0.21	0.16
p25	2.0	1.24	0.99
p75	5.0	39.48	24.49
p90	10.0	111.84	60.20

References

Akerlof, G. A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism*. *The Quarterly Journal of Economics*, 84(3):488–500.

Anand, A. and Venkataraman, K. (2016). Market conditions, fragility, and the economics of market making. *Journal of Financial Economics*, 121(2):327–349.

Angeris, G., Evans, A., and Chitra, T. (2020). When does the tail wag the dog? Curvature and market making. *arXiv:2012.08040 [q-fin]*. arXiv: 2012.08040.

Angeris, G., Evans, A., and Chitra, T. (2021). Replicating Market Makers. *arXiv preprint arXiv:2103.14769*.

Angeris, G., Kao, H.-T., Chiang, R., Noyes, C., and Chitra, T. (2019). An analysis of Uniswap markets. *arXiv preprint arXiv:1911.03380*.

Aoyagi, J. (2020). Liquidity Provision by Automated Market Makers. SSRN Scholarly Paper ID 3674178, Social Science Research Network, Rochester, NY.

Aquilina, M., Budish, E., and O’Neill, P. (2021). Quantifying the High-Frequency Trading “Arms Race”. *The Quarterly Journal of Economics*, (qjab032).

Aspris, A., Foley, S., Svec, J., and Wang, L. (2021). Decentralized exchanges: The “wild west” of cryptocurrency trading. *International Review of Financial Analysis*, page 101845.

Biais, B., Foucault, T., and Moinas, S. (2015). Equilibrium fast trading. *Journal of Financial Economics*, 116(2):292–313.

Black, F. (1971a). Toward a fully automated stock exchange, part I. *Financial Analysts Journal*, 27(4):28–35. Publisher: Taylor & Francis.

Black, F. (1971b). Toward a fully automated stock exchange, Part II. *Financial Analysts Journal*, 27(6):24–28. Publisher: Taylor & Francis.

Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A., and Sokolov, K. (2016). High-Frequency Trading and Extreme Price Movements. SSRN Scholarly Paper ID 2531122, Social Science Research Network, Rochester, NY.

Brogaard, J., Hagströmer, B., Nordén, L., and Riordan, R. (2015). Trading Fast and Slow: Colocation and Liquidity. *Review of Financial Studies*, page hhv045.

Brogaard, J., Nguyen, T. H., Putniņš, T. J., and Wu, E. (2021). What Moves Stock Prices? The Role of News, Noise, and Information. SSRN Scholarly Paper ID 3105862, Social Science Research Network, Rochester, NY.

Budish, E., Cramton, P., and Shim, J. (2015). The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response. *The Quarterly Journal of Economics*, 130(4):1547–1621.

Budish, E. B., Lee, R. S., and Shim, J. J. (2020). A Theory of Stock Exchange Competition and Innovation: Will the Market Fix the Market? SSRN Scholarly Paper ID 3391461, Social Science Research Network, Rochester, NY.

Capponi, A. and Jia, R. (2021). The Adoption of Blockchain-based Decentralized Exchanges. SSRN Scholarly Paper ID 3805095, Social Science Research Network, Rochester, NY.

Chaboud, A., Chiquoine, B., Hjalmarsson, E., and Vega, C. (2014). Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. *The Journal of Finance*, 69(5):2045–2084.

Daian, P., Goldfeder, S., Kell, T., Li, Y., Zhao, X., Bentov, I., Breidenbach, L., and Juels, A. (2019). Flash Boys 2.0: Frontrunning, Transaction Reordering, and Consensus Instability in Decentralized Exchanges. *arXiv:1904.05234 [cs]*. arXiv: 1904.05234.

DODO (2020). DODO – A Next-Generation On-Chain Liquidity Provider Powered by Pro-active Market Maker Algorithm.

Dyhrberg, A. H., Foley, S., and Svec, J. (2019). When bigger is better: The impact of a tiny tick size on undercutting behavior. SSRN Scholarly Paper ID 3194932, Social Science Research Network, Rochester, NY.

Easley, D., López de Prado, M. M., and O'Hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. *The Review of Financial Studies*, 25(5):1457–1493.

Foucault, T. (1999). Order flow composition and trading costs in a dynamic limit order market. *Journal of Financial Markets*, 2(2):99–134.

Foucault, T., Kozhan, R., and Tham, W. W. (2016). Toxic Arbitrage. *The Review of Financial Studies*, 30(4):1053–1094.

Foucault, T., Röell, A., and Sandas, P. (2003). Market making with costly monitoring: An analysis of the SOES controversy. *The Review of Financial Studies*, 16(2):345–384.

Glosten, L. R. and Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100.

Griffin, J. M. and Shams, A. (2020). Is Bitcoin Really Untethered? *The Journal of Finance*, 75(4):1913–1964. onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12903.

Han, J., Huang, S., and Zhong, Z. (2021). Trust in DeFi: An Empirical Study of the Decentralized Exchange. SSRN Scholarly Paper ID 3896461, Social Science Research Network, Rochester, NY.

Hendershott, T., Jones, C. M., and Menkveld, A. J. (2011). Does Algorithmic Trading Improve Liquidity? *The Journal of Finance*, 66(1):1–33.

Hendershott, T. and Menkveld, A. J. (2014). Price pressures. *Journal of Financial Economics*, 114(3):405–423.

Ho, T. and Stoll, H. R. (1980). On dealer markets under competition. *The Journal of Finance*, 35(2):259–267.

Ho, T. and Stoll, H. R. (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics*, 9(1):47–73.

Hoffmann, P. (2014). A dynamic limit order market with fast and slow traders. *Journal of Financial Economics*, 113(1):156–169.

Kirilenko, A., Kyle, A. S., Samadi, M., and Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, 72(3):967–998.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica*, 53(6):1315–1335.

Lehar, A. and Parlour, C. A. (2021). Decentralized Exchanges.

Malceniece, L., Malcenieks, K., and Putniņš, T. J. (2019). High frequency trading and comovement in financial markets. *Journal of Financial Economics*, 134(2):381–399.

Menkveld, A. J. (2013). High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4):712–740.

Park, A. (2021). The Conceptual Flaws of Constant Product Automated Market Making. SSRN Scholarly Paper ID 3805750, Social Science Research Network, Rochester, NY.

Pastor, L. and Stambaugh, R. F. (2019). Liquidity Risk After 20 Years. Working Paper 25774, National Bureau of Economic Research. Series: Working Paper Series.

SEC (2019). Division of Trading and Markets: Background Paper on the Market Structure for Thinly Traded Securities. Staff Background Paper, US Securities and Exchange Commission.

Shkilko, A. and Sokolov, K. (2020). Every cloud has a silver lining: Fast trading, microwave connectivity, and trading costs. *The Journal of Finance*, 75(6):2899–2927. Publisher: Wiley Online Library.

Stoll, H. R. (1978). The supply of dealer services in securities markets. *The Journal of Finance*, 33(4):1133–1151. Publisher: Wiley Online Library.

Stoll, H. R. (1989). Inferring the components of the bid-ask spread: Theory and empirical tests. *The Journal of Finance*, 44(1):115–134.

van Kervel, V. (2015). Competition for Order Flow with Fast and Slow Traders. *The Review of Financial Studies*, 28(7):2094–2127.

Van Kervel, V. and Menkveld, A. J. (2019). High-frequency trading around large institutional orders. *The Journal of Finance*, 74(3):1091–1137.

Wang, Y., Heimbach, L., and Wattenhofer, R. (2021). Behavior of Liquidity Providers in Decentralized Exchanges. *arXiv:2105.13822 [q-fin]*. arXiv: 2105.13822.

Weller, B. M. (2018). Does Algorithmic Trading Reduce Information Acquisition? *The Review of Financial Studies*, 31(6):2184–2226.

Xu, J., Vavryk, N., Paruch, K., and Cousaert, S. (2021). SoK: Decentralized Exchanges (DEX) with Automated Market Maker (AMM) protocols. *arXiv preprint arXiv:2103.12732*.