The impact of ETF index inclusion on stock prices

John Duffy Daniel Friedman Jean Paul Rabanal
Olga A. Rud*

October 25, 2023

Abstract

We report on a laboratory experiment examining how demand for ETF index products affects the prices and trading volume of assets. We compare an environment where the ETF index includes all assets in a market against an environment where a redundant asset is excluded from the index. We find that (i) subjects place significant value on the ETF index asset beyond the value of its constituent assets; (ii) there is a substantial index premium for included assets; and (iii) the index premium persists even when short-selling is permitted. The price increases of the constituent assets and of the ETF itself suggest that ETF products can distort markets to some degree.

Keywords: index inclusion premium, ETFs, experimental finance

*John Duffy (duffy@uci.edu) is at University of California, Irvine; Daniel Friedman (dan@ucsc.edu) is at University of California Santa Cruz, and University of Essex; Jean Paul Rabanal (jeanpaulrab@gmail.com) and Olga Rud (olga.rud@gmail.com) are at the University of Stavanger. We are grateful for the comments received by the Editor, the Associate Editor, two anonymous referees, Peter Bossaerts, Stephen Brown, Doug Davis, Nobuyuki Hanaki, Debrah Meloso, Bernt-Arne Ødegaard as well as from participants in several seminars and workshops. We also gratefully acknowledge research support for this project from the IFREE Small Grants program. Morgan Grant provided invaluable assistance in developing the laboratory software. The pre-registration material can be found at https://www.socialscienceregistry.org/trials/8498

1 Introduction

The index inclusion premium refers to the increase in the price of a stock following its inclusion in a market index. Empirical literature has documented the index inclusion premium, and investigated several explanations for it, mainly through *natural experiments* in the field. For instance, Chang et al. (2015) exploit a mechanical index reconstitution event and use regression discontinuity methods to measure the mean index premium for new additions to the Russell index.[1] Other studies focus on unique events to study cross-sectional returns (e.g. Greenwood, 2005 and Kaul et al., 2000) and found positive (negative) abnormal returns for new index additions (deletions).

Shleifer (1986) proposed that the index inclusion premium arises from downward sloping demand curves for included assets. If the demand for such assets is downward sloping (e.g., because there are no perfect substitutes) rather than horizontal (as is assumed in neoclassical finance), then inclusion of the asset in the index shifts the demand curve for that asset to the right and, assuming the quantity of the asset does not change, drives the price of the included asset up. Shleifer (1986) provided evidence for this explanation for the index premium associated with stocks added to the S&P 500 index.[2]

It is nevertheless unclear from the current literature whether the downward sloping demand explanation stems from demand shocks (i.e., investors' demand for an index which propagates to constituent assets), or is instead due to the enhanced liquidity of assets included in an index (i.e., investors' preference to trade for more liquid assets creating some network externalities). Furthermore, it is not well understood to what extent limits to arbitrage are needed to observe an index inclusion premium. For example, Wurgler and Zhuravskaya (2002) and Greenwood (2005) argue that arbitrage capital is insufficient to offset the demand shocks that increase the value of included index stocks compared with the excluded stocks.

In this paper, we propose a laboratory experiment to uncover the mechanisms behind the index premium. We also consider whether relaxation of limits to arbitrage, in particular restrictions on short-selling, are important for sustaining such a premium.

[1] Ben-David et al. (2018) also adopt such techniques to identify the impact of ETF index assets on the constituent assets regarding volatility.

[2] Another possible explanation for the index premium is a signalling effect, wherein inclusion signals future growth in the stock price (Denis et al., 2003, Dhillon and Johnson, 1991 and Jain, 1987), but this does not seem to be the mechanism that drives the index premium in Chang et al. (2015) or Greenwood (2005).

A laboratory environment offers the advantage of direct control over the fundamental values of all assets, and therefore allows us to study the impact of index inclusion, not only on recently included (or recently excluded) assets, but also on assets that were always included in the index. That is, we are able to detail in a simple setting how index inclusion affects an entire asset market.

In modern markets, ETF assets offer the advantage of trading an index in both a primary market and in a secondary market. The primary market is open only to Authorized Participants (APs) who can create and redeem ETF shares. The secondary market, which accounts for about 90% of overall trading volume, is open to all investors (ICI, 2021). Our laboratory environment relegates the role of APs to computer algorithms ("bots") that enforce the law of one price between the ETF and the constituent assets when ETFs are created or redeemed, much as APs do in the field.

Our laboratory market allows human subjects to trade in three risky assets, denoted A, B and C, whose final payoff is determined by the state of nature. Asset C is identical to asset B in all states of nature, i.e., B and C are perfectly positively correlated.

Different experimental treatments have different ETF constituent assets but without altering the market portfolio. The baseline treatment, which we refer to as treatment *ABC*, includes an ETF (asset D) which is constructed using the market portfolio, i.e., one unit of each risky asset. The second environment, treatment *A2C*, excludes asset B from the ETF and replaces it with a second unit of the payoff-identical asset C. We define the ETF index premium as the difference-in-difference in the price of assets B and C across environments *A2C* and *ABC*. The third and final treatment, $A2C_{short}$, allows subjects to short-sell the constituent assets. This treatment allows us to investigate how the ETF index premium responds to relaxing one potential limit to arbitrage, the ability to short assets.

In our experiment, the creation of the index asset (ETF) is endogenous. Initially, the market has a zero supply of ETFs. The AP bot offers a unit of the ETF for sale using the best ask prices for the constituent assets. In treatment *ABC*, the ETF product is only generated when there are outstanding asks for all three constituent assets A, B, and C. In treatments *A2C* and $A2C_{short}$, the ETF is only generated when there are outstanding asks for at least one unit of A and two units of C. In that case, the bot posts an ETF ask price equal to the sum of the constituent asks. If a subject enters a bid for the ETF that crosses the bot's ask, that subject buys a new unit of the ETF as in a primary market. That ETF unit can then be traded in the secondary market.

Upon selling a new unit of the ETF, the AP-bot immediately offsets its position (with a latency of about 5 to 10 milliseconds) by submitting market bid orders for the ETF's constituent assets. The AP-bot holds the constituent assets in a trust out of which it pays dividends to the ETF owners. In that sense, the bot maintains a zero net supply. This procedure virtually eliminates possible losses for the bot; it can earn a modest profit when pre-existing bids for the ETF are higher than the bot's asks.[3]

Thus, the key distinctive features of our laboratory environment include: (i) perfect substitutability between some included and excluded assets; (ii) only AP-bots and profit-motivated subjects (there are no specialized participants such as passive index funds); (iii) an exploration of limits to arbitrage explanations via relaxation of short-sales constraints; and (iv) detailed analysis of complete market performance, including bid/ask spreads, order imbalances, transaction prices and trade volumes, and subjects' final risky asset holdings, for both excluded and included assets as well as for always-included assets.

Our results show evidence of a substantial index premium, despite feature (i); downward sloping demand curves for assets despite feature (ii); reduced and/or similar asset prices and spreads, and the persistence of ETF inclusion premia when limits to arbitrage are relaxed via feature (iii); and finally, a strong preference for diversification at minimum cost as the main mechanism driving the observed inclusion premium. Indeed, an advantage of our laboratory approach is that we are able to consider and rule out other possible mechanisms for the inclusion premium. Specifically, we rule out signalling explanations by design. We do not find evidence for a liquidity driven mechanism consistent with Pagano (1989), since we observe that the turnover for the included asset is not altered across treatments while its bid-ask spread increases. Finally, by relaxing short-selling constraints, we do not find evidence for a limits to arbitrage explanation (see, e.g., Wurgler and Zhuravskaya, 2002, and Greenwood, 2005).

The persistence of the index premium after lifting short-selling constraints suggests that the index premium is mostly driven by the price impact of the creation of new ETF units. We observe that ETF prices of new units are higher than the NAV (Net Asset Value, the current total value of constituent assets) in the secondary market. We also find that the spread between the bid for the included asset C and the ask

[3] Appendix D shows that AP-bot profits are actually very small; most primary market transactions occur when the ETF price is equal to the NAV, defined as the sum of the prices for the constituent assets.

for the excluded asset B is negative or close to zero, suggesting that there are seldom profitable opportunities for the strategy *buy B & sell C.*

The experimental finance literature (e.g., Bossaerts and Plott, 2004) has consistently found that prices converge towards their equilibrium values, even in difficult scenarios (Asparouhova et al., 2015). Our study extends that work by introducing an ETF index asset. We find that demand for the ETF asset leads to increased prices for included assets, which in turn increases the price of the ETF. Thus, we find that index inclusion/exclusion can distort market prices, due to a desire to hold the market portfolio. To the best of our knowledge, our study is the first to document the complete impact of ETFs on a financial market, albeit a very simple one.

There is little work in the experimental literature on the consequences of index products. Duffy et al. (2021) were the first to study ETF index products in the laboratory and found that an ETF index asset can provide a useful asset pricing benchmark and reduce the mispricing of the ETF's constituent assets, when the ETF index covers the entire market. We confirm these findings in our ABC treatment: when the index covers the entire market, stocks are priced according to fundamentals. However, Duffy et al. (2021) do not consider the case of index reconstitution nor do they allow for the creation of new ETF units by APs in the primary market, which has important consequences for asset pricing.[4] Our paper is also related to the literature on the law of one price (LOP) in asset markets. Fisher and Kelly (2000), Childs and Mestelman (2006), and Charness and Neugebauer (2019) find that the LOP holds for two assets that have perfectly positively correlated returns. We present a case where two assets having the exact same fundamental value may be priced differently depending on whether or not they are included in an ETF index asset, in violation of the LOP, and we provide insights into the mechanisms responsible for this finding.

2 Environment

There are N subjects who trade 3 risky assets, $\theta \in \{A, B, C\}$, and an index ETF asset, whose composition is known to all subjects: (1) $ETF = A + B + C$ in the ABC treatment, or (2) $ETF = A + 2C$ in the $A2C$ treatment. All subjects are endowed with some combination of risky assets θ, as described in Table 1, and an interest-free loan L of cash. The market supply of each asset, θ, is 2 units per capita, and the ETF

[4] For an overview of experiments in a multiple asset environments, see, e.g., Duffy et al. (2022).

is constrained to be in zero net supply. A third treatment, $A2C_{short}$, is the same as treatment $A2C$ except that subjects in $A2C_{short}$ can short sell each of the θ assets.

Table 1: Endowment per subject type

Type	A	B	C	ETF	Loan (L)
I	3	1	1	0	210
II	1	3	3	0	210
Per capita	2	2	2	0	210

Note: There are 2 types of subjects: type I and type II. Regardless of type, each subject type receives a loan equal to 210 units of cash.

The payoff to each asset in each period depends on the state of nature $s \in$ {X, Y, Z}, where each state is equiprobable, as shown in Table 2. These state-contingent payoffs are publicly announced, but the realized state s is not known until the end of each trading period. The difference in expected payoffs between A (80), and B and C (60) is intended to highlight that assets B and C are identical but asset A is different. Note that the fundamental (state-contingent) value of the ETF index asset does not vary across treatments because B and C yield the same terminal payoffs. Since the ETF is offered at zero net supply, we can use the standard asset pricing model to predict asset prices.

Table 2: Asset payoff

Asset\State	X	Y	Z	Expected value
A	0	120	120	80
B	90	0	90	60
C	90	0	90	60
ETF	180	120	300	200
Probability	1/3	1/3	1/3	–

Note: The ETF payoff is equal across treatments because $B = C$, and the ETF composition is either $A + B + C$ or $A + 2C$.

At the end of each trading period, subjects' payoffs are computed by summing the payoffs of their portfolios of asset holdings at the realized state s, plus their final cash holdings minus their initial (loaned) cash. In treatments allowing short-selling, the value of a final short position in an asset (number of units shorted times state-contingent payoff) enters negatively. Of course, transient short positions can be covered by subjects within each trading period.

To determine the equilibrium price of assets θ, we assume mean-variance utility $U = \mu - \frac{b}{2}\sigma^2$. According to the CAPM model (see, e.g., Hirshleifer and Riley, 1992 and Appendix A for details), the price is

$$p^*_{\theta} = \mu - b\Delta\bar{Q}, \tag{1}$$

where μ is a vector with the expected value of assets θ, Δ is the covariance matrix of risky assets, and $\bar{Q}$ is the total supply of assets per capita. Assuming a moderate degree of risk-aversion, $b = 0.001$, leads to $p^*_A = 78$ and $p^*_B = p^*_C = 55$.

Assets are traded via a continuous double auction (CDA). Subjects can trade in up to four separate markets simultaneously, with one market for each asset. In the CDA format subjects participate in any market by submitting a limit order to buy (a bid) or to sell (an ask) a single unit, or by accepting an existing bid or ask. Existing limit orders for all four markets are displayed in each market in descending order for bids, from highest to lowest, and ascending order for asks, from lowest to highest. In case of a tie in order prices, the first order submitted has priority.

A market transaction occurs whenever a bid (b) and an ask (a) overlap, such that $b_\theta \geq a_\theta$, or when a subject accepts an existing order by clicking on it and agreeing to the bid or ask price. The transaction price is always that of the earlier-in-time order. As previously noted, in treatment $A2C_{short}$, subjects are allowed to short-sell each of the three risky assets θ, but not the ETF.

2.1 Authorized participants

In existing ETF markets, the creation and redemption of ETF shares in the primary market is handled by *Authorized Participants, AP* (typically banks or financial institutions) that adjust the number of outstanding ETF shares to keep the price of the ETF aligned with the value of its constituent assets. In our setting, the APs are robot or "bot" players. Our AP-bot is programmed to eliminate arbitrage opportunities in the market. Specifically, they ensure that the price of the ETF is equal to its net asset value (NAV) when creating or destroying ETF units, such that $P_{ETF} = P_A + P_B + P_C$ in the ABC treatment and $P_{ETF} = P_A + P_C + P_C$ in the $A2C$ and $A2C_{short}$ treatments. If these conditions do not hold, then subjects can profit from arbitrage opportunities.

Our AP-bot is tasked with the creation and redemption of ETF shares, while maintaining a zero net supply of ETFs. The AP-bot can offer the ETF for sale only if there exists outstanding asks for all the constituent assets. For example, in treatment ABC, if there are outstanding ask orders for assets A, B, C, denoted as a_A, a_B and a_C, then the AP will post an ask order for the ETF asset such that in the ABC treatment $a_{ETF} = a_A + a_B + a_C$. If a subject purchases the ETF offered by the AP-bot, then the bot will immediately accept the outstanding best asks for the assets A, B, and C and place these assets in a trust account out of which it will pay dividends to the ETF holders. Similarly, if the AP-bot finds outstanding bid orders for all of the constituent assets, then it will post a bid order for the ETF asset such that in the ABC treatment, $b_{ETF} = b_A + b_B + b_C$. If a subject accepts that bid, then the AP-bot will immediately accept the bids for the constituent assets releasing those assets from its trust account to the successful bidders.

2.2 Hypotheses

In this section we propose testable hypotheses that focus on: (i) the existence of an index premium, (ii) order imbalance, (iii) the impact of a non-binding short-selling constraint, and (iv) potential transmission mechanisms. Our motivation for these proposed hypotheses is derived from theoretical literature, primarily Greenwood (2005) and Pagano (1989), as well as some empirical and experimental work.

Hypothesis 1:

(a) *In treatment* A2C*, the price of asset C (B) will be greater (smaller) than in the baseline treatment* ABC*, consistent with a positive ETF index inclusion premium.*

(b) *A positive ETF index inclusion premium is associated with a larger order imbalance for the included asset compared to the excluded asset from the ETF index.*

Hypothesis 1(a) is derived from the empirical literature, e.g., Chang et al. (2015), who report that stocks that are added to an index experience price increases, while stocks that are deleted from an index experience price declines. However, H1(a) is contrary to the limits to arbitrage model of Greenwood (2005) in which "stocks whose fundamentals are positively correlated with stocks experiencing positive demand shocks will experience increases in price...even though no direct change in demand has occurred." In our experimental design, B and C are perfectly correlated, and thus if the price of C increases, so should the price of B.

To test for the ETF index premium of Hypothesis 1(a) we compare the price difference between equivalent assets B and C in treatment $A2C$ against the price difference in treatment ABC. The price difference between the two assets in a given treatment

M can be written as $\varphi_P^M := P_C - P_B$. Thus, we define the ETF index (I) premium as

$$\varphi_I := \varphi_P^{A2C} - \varphi_P^{ABC}. \quad (2)$$

Given that there are no differences in their state-contingent payoffs, any price difference between assets B and C can only be attributed to an ETF index inclusion. In treatment *A2C* we hypothesize that the price of asset C (B) is larger (smaller) than in treatment *ABC*. It follows that there will be a positive ETF index premium $\varphi_I > 0$ in equation (2).

Consistent with the ideas of Greenwood (2005), there should be no systematic difference in the demand for asset A across treatments, since A is always included in the index, and thus the price of A should remain the same in all treatments.

Hypothesis 1(b) is also derived from the empirical literature e.g., Chordia and Subrahmanyam (2004) who find that changes in asset prices are often presaged by order imbalances that reflect latent excess demand for an asset. To test this notion, we propose a measure of order imbalance, where $d \in (0, 1]$ is a weight parameter that discounts orders more heavily the further they are away from the midpoint $m(t)$ between the current best bid and ask prices, and $Q(p, t) \in Z$ is the number of outstanding + sell (− buy) orders at price p at time t. Following the existing literature, we then define the order imbalance as

$$z(t, d) := -\sum_p Q(p, t) d^{|p-m(t)|}. \quad (3)$$

In our data analysis we use $d = 0.99$ to give almost full weight to all serious orders and much lower weight only to extreme outliers.

Our hypothesis concerns the difference in the order imbalance (z) for the identical assets B and C in a treatment M as $\varphi_z^M := z_C - z_B$, and the difference in order imbalance difference (O) across treatments,

$$\varphi_O := \varphi_z^{A2C} - \varphi_z^{ABC}. \quad (4)$$

Preferences for index products and the constituent assets should lead to an increase in the demand for C, and a decrease in the demand for B in treatment *A2C*. The increase in demand should be related the number of bids relative to asks in the order book. For C (B), we expect to observe more (less) bids relative to asks. Thus, we expect that $z_B < z_C$, which results in $\varphi_z^{A2C} > 0$. In treatment *ABC*, where B and C are both included in the index, we do not expect a significant value for φ_z^{ABC}. Therefore, we expect that $\varphi_O > 0$ in equation (4).

Hypothesis 2: *Short-selling decreases the index premium, and the order imbalance.*

This hypothesis derives from the empirical literature which has rationalized the inclusion premium by assuming that there are limits-to-arbitrage. For example, Wurgler and Zhuravskaya (2002) and Greenwood (2005) argue that arbitrage capital is insufficient to offset the demand shocks that increase the value of the included stocks compared with the excluded stocks. That argument suggests that relaxing those limits to arbitrage, as in our treatment $A2C_{short}$, will reduce the order imbalance and the inclusion premium, consistent with Hypothesis 2.

Indeed, the limits-to-arbitrage theory might be read to suggest that when short-selling constraints are non-binding, the inclusion premium will vanish. Our Hypothesis 2 does not go that far, recognizing the possibility that the demand for the constituent assets may remain downward sloping (e.g. see Chang et al., 2015 or Greenwood, 2005). Furthermore, previous experimental work (Haruvy and Noussair, 2006) finds that while prices are lower when short-selling constraints are relaxed, they still fail to track the fundamental values. Hence Hypothesis 2 predicts only that the ETF index premium and order imbalance follow $\varphi_I^{short} < \varphi_I$ and $\varphi_O^{short} < \varphi_O$.

Hypothesis 3: *(a) Subjects will trade away from initial allocations so that their final risky asset allocations are closer to the market portfolio. Holding the ETF index asset will provide a convenient (low transactions cost) way to obtain the market portfolio. (b) Strong demand for the ETF will push its price above NAV in the secondary market, and will drive its order imbalance above that of the other risky assets.*

H3(a) follows from the standard financial asset pricing model (e.g., Sharpe, 1964) which suggests that investors should hold the market portfolio. In all treatments, the ETF provides the market portfolio, and thus purchasing it facilitates diversification.

The ETF asset provides a convenient vehicle for moving towards the market portfolio. Buying the ETF index from the AP-bot or from another subject also provides immediacy: a trader instantaneously acquires a unit of the market portfolio at a known price. By contrast, piecemeal purchase of the constituent assets in a CDA involves a time interval over which the trader is unsure of the price (or even availability) of the remaining pieces of the market portfolio unit. Of course, some human subjects will care more about diversification than others. Consequently, we expect that subjects with larger final ETF holdings will achieve a final allocation that is closer to the market portfolio as compared to subjects who do not hold ETF shares.

H3(b) predicts that the demand for the ETF will show up as a larger positive order imbalance, and will ultimately result in ETF prices in the secondary market that are above the NAV of the ETF. Recall that in the primary market, the ETF is created such that the ETF price follows the NAV. However, it is not necessarily the case that this holds when we look at transactions in the secondary market between human subjects.

Hypothesis 4: *The liquidity derived from the demand for ETFs will result in higher turnover and lower spreads for included assets.*

This hypothesis is derived from the market microstructure literature on liquidity and asset prices. For example, Pagano (1989) suggests that the liquidity of an asset is associated with higher trade volume and greater absorptive capacity. When an asset is more liquid, it is easier and cheaper to trade, which attracts more traders and increases trading volume. Similarly, smaller (larger) bid-ask spreads on assets also indicate greater (lower) liquidity of those assets see e.g., Foucault et al. (2013). In our context, strong demand for ETF products may lead to higher trade volume and lower bid-ask spreads for the included assets as compared to the excluded assets.

3 Laboratory procedures

The experiment was computerized using oTree (Chen et al., 2016) and conducted using subjects from the Experimental Social Science Laboratory at the University of California, Irvine. Subjects included undergraduate students from all fields of study. In all, we recruited 182 subjects to participate. Each subject participated in a single session of one of the three treatments: $\{ABC, A2C, A2C_{short}\}$. In each session, subjects were given written instructions which were read aloud. They were then asked to complete a comprehension quiz to check their understanding of the written instructions. Copies of the instructions and quiz questions can be found in the Appendix B. Upon completion of the quiz, subjects received feedback as to which quiz questions they answered correctly or incorrectly and in the latter case, they were instructed about the correct answer. The experimenter (one of the authors) answered any remaining questions privately.

Each session lasted just under 90 minutes and included 10 market trading periods, with 3 practice periods at 3 minutes each, and 7 real periods at 5 minutes each. One of the last 7 periods was randomly selected for payment at the conclusion of the experiment. Subjects' point totals from the chosen period were converted into dollars at the rate of $4 per 100 points. Following each period, subjects received feedback regarding the value of their portfolio holdings (which depended on the realized state), their remaining cash on hand less the value of the loan, all converted into points, and thus their final point earnings in that period. As summarized in Table 3, we conducted 15 sessions, 5 for each of our 3 treatments, with between 10 and 16 subjects per session. Average earnings are $13.95 which excludes a show-up fee of $7.[5]

Table 3: Overview of experimental sessions

Treatment	Sessions	Subjects	Earnings ($, mean)
ABC	5	58	12.81
$A2C$	5	62	16.40
$A2C_{short}$	5	62	12.57
Total	15	182	13.95

Note: Each session had 10-16 subjects. Each subject also received $7 as show-up fee.

3.1 Trading interface

Figure 1 shows the market user interface (UI). The top panel of the UI is divided into four quadrants, where each quadrant represents an asset market. In each asset market, a subject can view the order book with outstanding Bids and Asks, and all past traded prices (Trades). The bottom panel of the UI also has four parts and shows (i) current portfolio holdings, (ii) the composition of the ETF index asset (labelled asset D for a neutral framing), (iii) the time remaining for trade in the period, and (iv) terminal asset payoffs according to the three possible states of nature.

Subjects can submit single unit limit orders by typing a buy (sell) price in the relevant input box for that asset, and clicking on the button labeled "Buy" ("Sell"). A subject's submitted order is immediately posted to the order book for the relevant

[5]Higher earnings in treatment $A2C$ are due to more frequent realizations of state Z, which yields a higher overall payoffs.

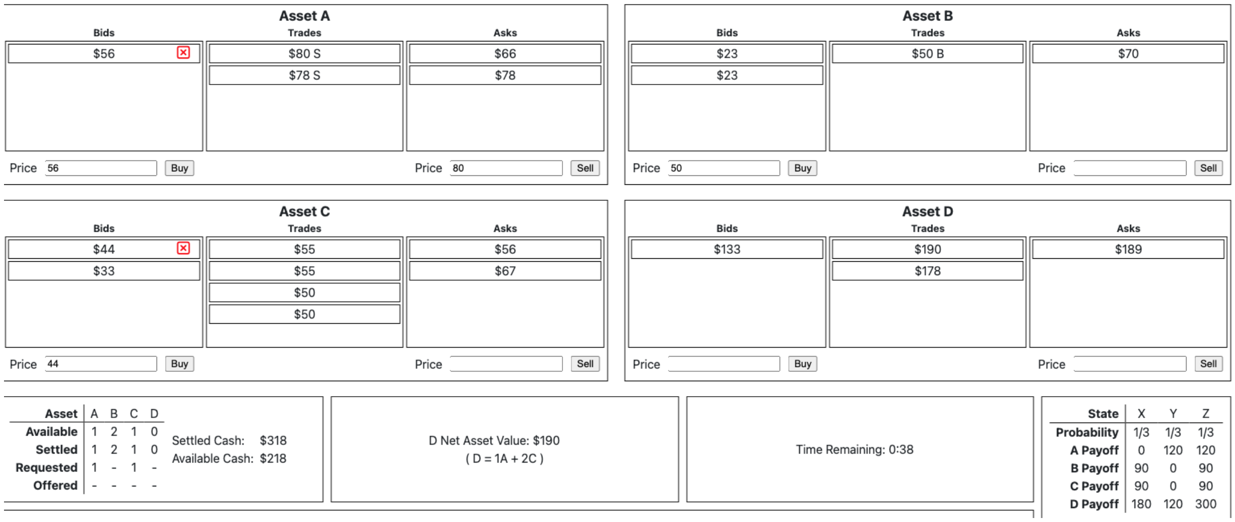

Figure 1: User interface (UI) $A2C_{short}$ treatment; other interfaces are similar.
Note: A red cross appears beside the subject's own orders; by clicking it she can cancel the order. Her own trades are marked by the letter S if she sold and by B if she bought. The AP-bot posts asks and bids for asset D (the ETF) using the current best bids and asks of the underlying assets, here a unit of A and two units of C.

asset and is identified on their screen by a red cross next to the order. The subject can cancel their own order if it is not yet executed by clicking on the red cross. Bids in the order book are sorted from highest to lowest and asks are sorted from lowest to highest.

Transactions occur in one of two ways. First, if a limit order bid (ask) comes in above (below) an existing limit order ask (bid) then a unit of the asset is traded at the existing limit order ask (bid) price. Second, market buy or sell orders can be made with immediate execution by clicking on any existing bids (asks) in any of the four asset market order books and clicking a "confirm" button. All transactions are recorded and presented in the center column of the order book under the heading "Trades". A letter "B" or "S" indicates that the transaction belongs to the subject viewing the book, and whether she bought (B) or sold (S) the asset unit. In all treatments, bids are rejected when they exceed the subject's cash balance. In the treatments ABC and $A2C$, the asset inventories cannot drop below zero (a no short-selling constraint), while in treatment $A2C_{short}$, the short-selling constraint is relaxed for θ assets, which are allowed to go below zero.[6] Here, if the final position in any asset is negative, then the subject must buy back the asset at its terminal value, as determined by the realized

[6] A subject can short-sell at most 100 units of each underlying θ asset. Given that the supply of such assets is two per capita, this limit will never bind. Of course, asset turnover is also constrained by available cash.

state, s. Asset D, the ETF asset, is a composite asset that is formed as a treatment-specific combination of the θ assets. At the start of each trading period, none of the subjects hold any D assets, but the AP-bot will offer units of asset D for sale as soon as other subjects enter asks orders for all its constituent assets. Subjects can submit bids for D at any point in the market, but they can submit asks for D only if they hold the ETF in their current portfolio.

4 Results

We begin our data analysis with an overview of asset prices and order imbalances from two representative sessions, one from treatment ABC as shown in Figure 2 and the other from treatment $A2C$ as shown in Figure 3. Similar figures for all sessions can be found in Appendix C.

In the top panel of Figure 2 we can see that the prices of B and C in the ABC treatment oscillate around their expected value (the dashed gray line at 60). The prices for the ETF and A are higher, and approach their expected values (the gray dashed lines at 200 and 80, respectively). The bottom panel in Figure 2 shows the order imbalance for all assets in the market. The order imbalances for the two identical assets B (red), and C (orange) fluctuate over similar ranges. These observations together suggest that assets B and C, which have the same payoff structure and are both part of the ETF index, are priced and traded similarly in the ABC treatment.

Figure 3 presents an example session from the $A2C$ treatment, where B is excluded from the index and is replaced by identical asset C. The top panel shows that the price of excluded asset B (red) is usually well below the price of included asset C (orange), despite the fact that the two assets are payoff-identical. Recall that we do not observe this price divergence in Figure 2, where both assets are included. These figures therefore suggest that being included (excluded) in the index increases (decreases) the asset price, i.e., that there is an ETF index premium.

To explore the forces behind the apparent price divergence, we compare the order imbalances. The bottom panel of Figure 3 shows that the order imbalance for B (red line) is consistently below the order imbalance for C (orange line) in the $A2C$ treatment. Indeed, the excluded asset (B) typically has a more negative order imbalance (i.e., there are more outstanding asks than bids), than the included asset (C). Thus, the idea of a positive price premium due to index inclusion is supported by an examination of

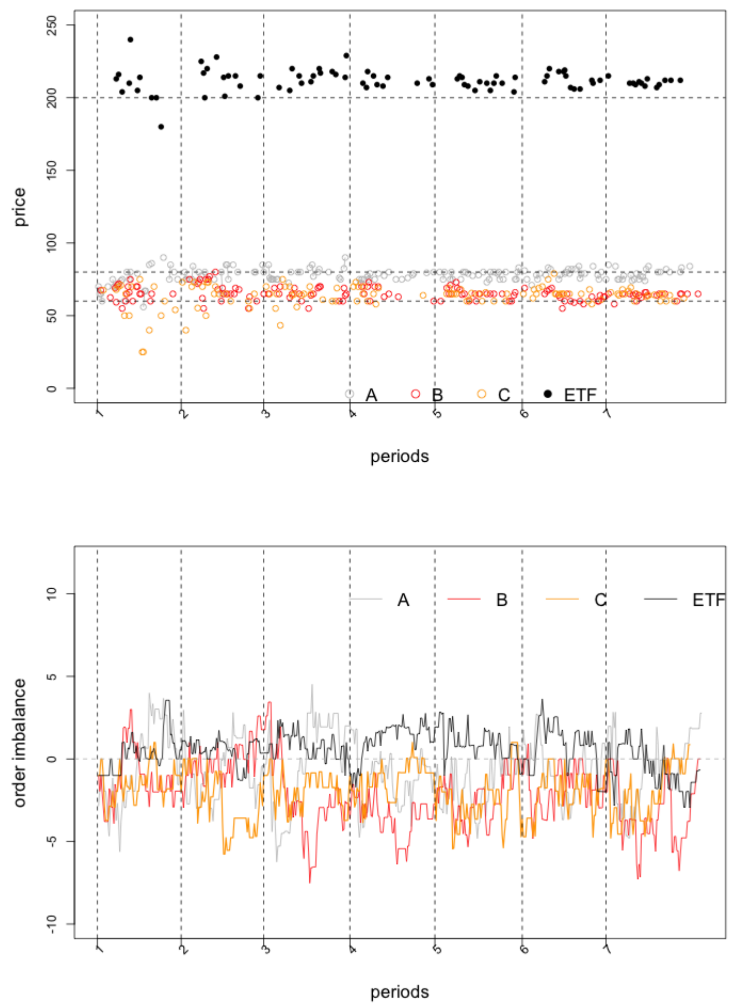

Figure 2: Prices and order imbalance for *ABC* session 3

Note: The top (bottom) panel presents asset prices (order imbalances in equation (3) with $d = 0.99$) for assets A, B, C, and the ETF. Each tick corresponds to four seconds and the dashed vertical lines denote the start of a trading period. In the top panel, the horizontal dashed lines represent the expected value of each asset: 200, 80 and 60 for the ETF, A, and B or C, respectively.

the order book, which shows a larger net demand for the included asset than for the excluded asset.

Figure 4 summarizes the period-by-period average price differences (left panel) and the average order imbalance differences (right panel) between assets B and C over all sessions of all three treatments. The black lines denote the *ABC* treatment, blue dashed lines the *A2C* treatment, and red dashed lines the $A2C_{short}$ treatment. Consistent with Hypothesis 1a, there is a price differential between the two identical assets B and C of about 20 in the *A2C* treatment and the $A2C_{short}$ treatment. In the *ABC* treatment,

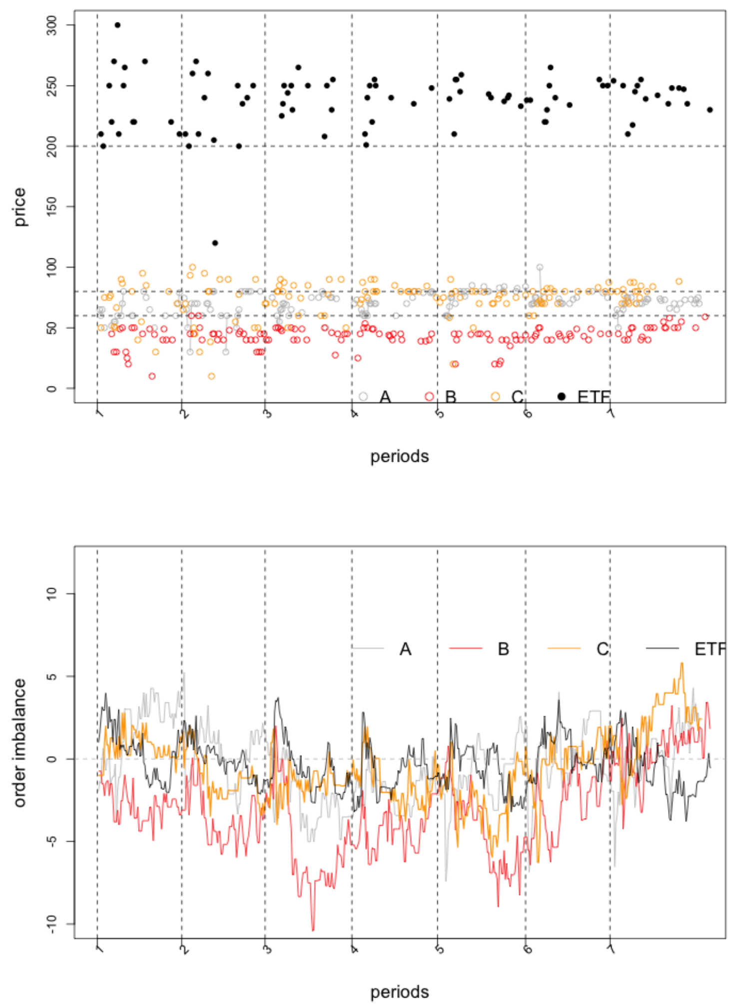

Figure 3: Prices and order imbalance for *A2C* treatment session 4

Note: The top (bottom) panel presents prices (order imbalances in equation (3)) for assets A, B, C, and the ETF. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. In the top panel, the horizontal dashed lines represent the expected value of the asset at 200, 80 and 60 for the ETF, A, and B or C, respectively.

however, the prices of the identical assets B and C appear to be the same; the price differential is close to zero. Also, consistent with Hypothesis 1b, the mean difference between the order imbalances for assets B and C of about 2 for treatments $A2C$ and $A2C_{short}$ is generally greater than the order imbalance observed in treatment ABC, where this value is usually close to zero. The greatest order imbalance is exhibited in treatment $A2C_{short}$, the treatment that permits short-selling.

Table 4 presents average prices and per-period turnover for each asset and treatment. The top panel includes all transactions, while the other two panels separate

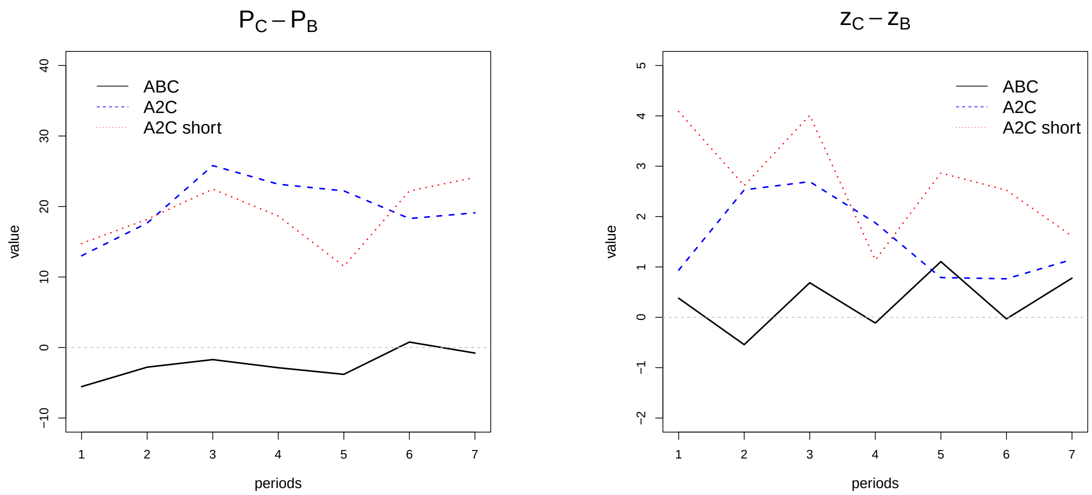

Figure 4: Mean difference in price (left) and order imbalance (right) for assets B and C by trading period and treatment across all sessions.

Note: Each session includes 7 periods, and mean differences over its 5 sessions are shown separately for each of the 3 treatments, ABC, $A2C$ and $A2C_{short}$.

transactions in the primary market (where ETFs are created) from those in the secondary market (no ETF creations). In our experiment, ETFs were never redeemed (destroyed); evidently AP-bot bids (equal to the sum of posted bids for the constituent assets) never exceeded subjects' bids for ETFs when trades occurred.

The price for A is very similar in treatments ABC and $A2C$. It is lower in $A2C_{short}$, likely due to the larger supply of assets when short-selling is permitted. The prices of B and C are similar in treatment ABC. However, those prices diverge in treatments $A2C$ and $A2C_{short}$; as noted earlier, the difference is about 20 in both treatments. Also note that the prices for the constituent assets appear to be higher in the primary market as compared to the secondary market. We discuss this more formally below.

The ETF price is near its expected value of 200 in treatment $A2C_{short}$ but is noticeably higher in the other treatments. Most of the ETF transactions are creations, averaging 5-7 units per period, versus only one unit transacted per period traded in the secondary market. Of course, primary market turnover for the constituent assets follows the index composition, while the constituent assets are all actively traded in the secondary market, each of them averaging around 7-13 transactions per period.

Table 5 presents the average order imbalance and spread of each asset per period for each treatment. We observe a negative order imbalance for all assets except for the ETF, which reflects the demand for index products. For the spread, we observe

Table 4: Mean price and turnover

	Price			Turnover		
	ABC	*A2C*	$A2C_{short}$	*ABC*	*A2C*	$A2C_{short}$
All transactions						
A	77	74	60	17	13	19
B	64	51	40	16	11	13
C	62	73	61	16	17	21
ETF	213	241	198	8	6	7
$\varphi_P^M = P_C - P_B$	-2	22	21	–	–	–
Primary market						
A	81	80	66	7	5	6
B	67	–	–	7	–	–
C	65	78	65	7	10	12
ETF	213	240	198	7	5	6
Secondary market						
A	74	71	57	11	8	13
B	62	51	40	10	11	13
C	61	66	55	10	7	9
ETF	206	248	201	1	1	1

Note: Mean price is the average price over all periods of all sessions for the given assets (rows) and treatments (columns). Turnover is similarly calculated for the number of units traded per period. Primary market transactions are for newly created ETFs, i.e., sales by AP-bots. Secondary market transactions are between subjects.

a value around 18-20 for the constituent assets in treatment *ABC* and around 41 for the ETF. The spread increases in treatment *A2C* to 30 for *A*, 24 for *B*, 46 for *C* and 100 for the ETF. Lifting the short-selling constraints decreases the spread of all assets relative to treatment *A2C*.

4.1 Test Results for Hypotheses 1 and 2

Result 1a: *Consistent with Hypothesis 1a, the ETF index premium* φ_I *is positive in treatment* A2C. *When an asset is excluded from the index it is priced lower than the identical included asset. The price of the always-included asset does not change between treatments.*

In top panel of Table 4, which uses data from all transactions, we can see that the price difference $\varphi_P^M := P_C - P_B$ is close to zero in treatment *ABC*, but is greater than zero in treatments *A2C* and $A2C_{short}$. More formally, we run a conservative non-parametric Fisher-Pitmann (FP) permutation test (Fisher, 1935, and Pitman, 1938),

Table 5: Order imbalance and spread

	Order Imbalance			Spread		
	ABC	*A2C*	$A2C_{short}$	*ABC*	*A2C*	$A2C_{short}$
A	-0.20	-0.67	-0.02	19.92	30.21	26.96
B	-2.02	-2.22	-4.28	17.69	23.63	15.54
C	-1.71	-0.68	-1.60	17.54	46.03	26.59
ETF	0.72	0.48	0.98	40.98	106.92	58.35
$\varphi_z^M = z_C - z_B$	0.31	1.54	2.68	–	–	–

Note: We compute the mean order imbalance in equation (3) and spread (best ask minus best bid) per asset and treatment.

using each session as a single observation of the price difference. It confirms Hypothesis 1a that there is a positive ETF index premium with a p-value of 0.007 for *A2C* and 0.01 for $A2C_{short}$). We conclude that when an identical asset is excluded from the index, it will be priced differently. In our data, the size of the ETF inclusion premium is large, about a third of the expected payoff for the excluded asset.[7]

Consistent with the ideas of Greenwood (2005), the FP test detects no significant change in the price of asset A, which is always included in the index, across treatments (p-values of 0.69 and 0.095 for treatments *A2C* and $A2C_{short}$). This is intuitive, since there is no systematic difference in the ETF demand between treatments.

Result 1b: *Consistent with Hypothesis 1b, the order imbalance becomes more positive for the identical included asset C in treatment* A2C*.*

This result is supported by the last row of Table 5. The order imbalance difference is small in treatment *ABC* (0.31) and becomes larger in treatment *A2C* (1.54); the FP test has a p-value of 0.04. This suggests that there is upward pressure on the price for C, as the number of asks relative to bids decreases.

Result 2: *Short-selling in treatment* A2C_{short} *does not decrease the ETF index premium observed in treatment* A2C*, contrary to Hypothesis 2. However, it does decrease order imbalances, prices, and spreads.*

We formally test whether the index premium differs across treatments *A2C* and $A2C_{short}$ by performing the non-parametric FP test. The results suggest that lifting

[7] In Appendix D, Table D.2 we look at the best asks and bids 30 seconds prior to market closing, and find that the best ask for C exceeds the best ask for B by 28. This spread in asks corroborates Result 1a.

short-selling constraints has no impact on the index premium (p-value of 0.874). This result may seem surprising at first, however there are seldom profitable opportunities to sell asset C and buy asset B; see Appendix D, Result 6.[8]

The prices of A and the ETF (Table 4) weakly decrease in treatment $A2C_{short}$ as compared with treatment $A2C$ (p-value of 0.079 and 0.085, respectively). The spreads for asset C and ETF in Table 5 (p-value of 0.098 and 0.083, respectively) also weakly decrease, while the spread for the excluded asset B significantly decreases (p-value of 0.043). The weak statistical significance is due in part to the conservative non-parametric approach we take in our analyses.

The order imbalance difference in $A2C_{short}$ (2.68) becomes larger compared to the baseline treatment (p-value of 0.009). The greater difference in the $A2C_{short}$ treatment is likely due to the larger number of constituent assets available for trade when the short-selling constraint is relaxed.

4.2 Underlying Mechanisms

Having documented the existence of an index premium, and the associated order imbalances, we now turn our attention to its source. We examine two potential mechanisms that could give rise to an index premium: (i) a strong preference for diversification at minimum cost, and (ii) enhanced liquidity arising from more active trading, and lower spreads.

Result 3a: *Consistent with Hypothesis 3a, most subjects' final asset positions are closer to the market portfolio than to their initial endowments. This is especially true for subjects who use a buy-and-hold ETF strategy.*

We capture departures from the market portfolio in the *portfolio imbalance statistic* $\frac{|2A-B-C|}{A+B+C+3D}$, where letters refer to median final holdings of units across trading periods by a particular human subject. The numerator measures the absolute distance to the market portfolio,[9] and the denominator counts the total number of constituent-equivalent asset units.

[8]We also observe the absence of arbitrage in the field. For example, short-term U.S. Treasury notes trade at a discount relative to U.S. Treasury bills of the same maturity date, despite having the same default risk. The reason, as Amihud and Mendelson (1991) point out, is the difference in liquidity and not the absence of short-sale opportunities.

[9]Recall from Table 1 that with n subjects, the market portfolio consists of exactly $2n$ units of each constituent asset, so a subject holding a fraction k of it will hold $(A, B, C) = 2nk(1, 1, 1)$, resulting in numerator $|2A - B - C| = 0$. Of course, the numerator (and thus the risky portfolio imbalance) is also 0 for the ETF in every treatment.

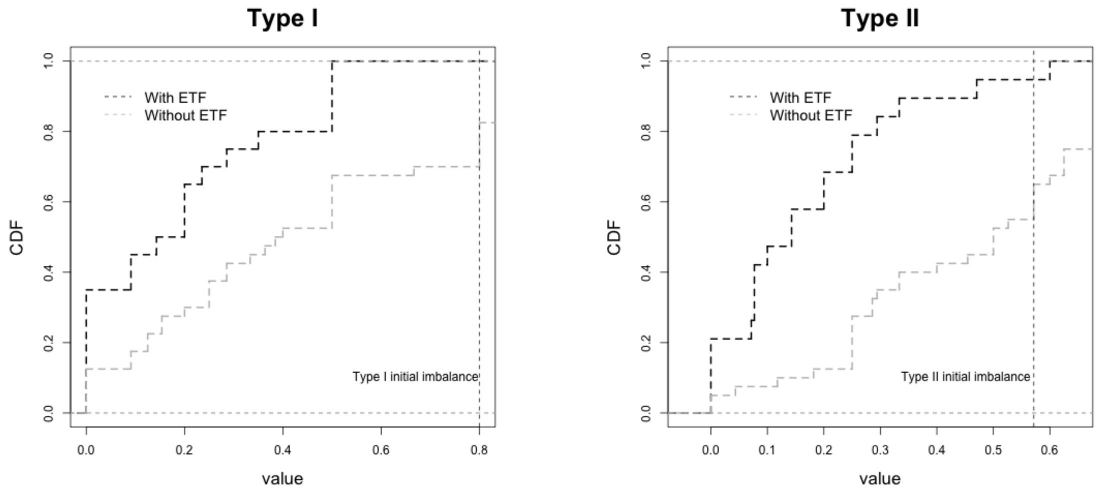

Figure 5: Subject's risky portfolio imbalance

Note: Subject's risky portfolio imbalance is computed as $\frac{|2A-B-C|}{A+B+C+3D}$ where the letters refer to the median final asset holdings. For the empirical cumulative distribution function, each observation represents a subject. Type I players (left) start with an imbalance of 4/5 while type II players (right) with an imbalance of 4/7. These values are represented by vertical dashed lines. We classify subjects with ETF holdings (black, 56 in total) and without ETF holdings (gray, 126 in total).

Figure 5 presents CDFs of final risky portfolio imbalances separately for subjects of endowment type I (left panel) and endowment type II (right panel) as defined in Table 1. We include all subjects in treatments *ABC* and *A2C*, but not those in $A2C_{short}$ because negative final holdings greatly complicate comparisons. We also separate subjects according to whether (black dashed line) or not (gray dashed line) they hold ETF units. Remarkably, all type I subjects who hold an ETF decrease their risky portfolio imbalance, and almost 40 percent of these subjects achieve a perfectly balanced portfolio. Type I subjects who do not hold the ETF asset exhibit a greater imbalance (p-value of 0.007 for the FP test with each subject as a single observation). For type II subjects, we also observe that the risky portfolio imbalance is greater for subjects who do not hold ETFs as compared with those who do (p-value < 0.001). While 90% percent of subjects who hold an ETF decrease their initial imbalance, less than 60% of those who do not hold an ETF are able to achieve such a reduction.

Table 4 shows that the majority of ETF transactions are driven by the creation of ETFs in the primary market. Although the AP-bot does post bids for the ETF asset, subjects are apparently unwilling to sell it to the bot. Thus, we do not observe any redemptions of ETFs. In the secondary market, we observe some ETF transactions between human subjects, but not very many. Indeed, as detailed in Appendix D, Table

D.1, only about 30% of ETF trades were in the secondary market in the *A2C* session with the most active secondary market. Thus, we can conclude that most subjects adopt a *buy-and-hold* strategy with regard to the ETF asset, buying it from the AP-bot and holding it for the duration of the market.

Result 3b: *Consistent with Hypothesis 3b, the preference for ETFs is reflected in (i) the ETF price in the primary market above the value of its constituents assets in the secondary market, and (ii) a positive order imbalance for the ETF, despite a non-positive order imbalance for all other assets.*

While previous lab experiments without ETFs (Bossaerts and Plott, 2004 and follow-up work) found that asset prices are generally below their expected value, suggesting a risk-premium, we find that the ETF share is typically priced at or above its expected value of 200. This may suggest that an ETF share, which offers the market portfolio, helps fulfill subject's diversification needs, in support of Hypothesis 3.

More formally, Table 4 shows that for treatment *ABC* the mean price of the ETF is 213, and we cannot reject the null that it is equal to the expected value of 200 (p-value of 0.231 using the FP non-parametric test). In treatment *A2C*, the mean ETF price increases by about 30, and we can reject that it is equal to the expected value of 200 (p-value of 0.006). In treatment $A2C_{short}$, where the short-selling constraint is relaxed, the mean ETF price is close to the expected value of 200 and we cannot reject the null of no difference (p-value of 0.954).[10]

We also analyze whether the demand for new ETFs pushes its price above the NAV in the secondary market. If so, this would suggest that there is a preference for the ETF bundle compared to its constituent assets. Using data from Table 4, we see that the mean NAV in the secondary market is 197, 203, and 167 for treatments *ABC*, *A2C* and $A2C_{short}$, respectively. These NAV values are all below the ETF prices observed in the primary market using the FP non-parametric test (p-value of 0.027, 0.005 and 0.019, respectively for each treatment).

Regarding order imbalances, Table 5 shows a positive value for the ETF asset (p-value of 0.001 for the non-parametric FP test using each session as a unique observation) while the rest of the assets have a non-positive value (p-values 0.296 for A, < 0.001 for

[10] We also test whether the ETF prices are equal between the *A2C* and *ABC* treatments. We find a $p-$value of 0.053 which suggests a statistically weak difference, originating from a greater demand for asset C when ETFs are issued, which in turn increases the ETF price.

B, and of 0.016 for C). A positive order imbalance means that the weighted sum of bid orders is larger than that of the ask orders, indicating upward pressure on prices.

Result 4: *The turnover of the included asset C does not change in treatment* A2C *compared to the baseline treatment ABC, while the bid-ask spread increases.*

Contrary to Hypothesis 4, we do not find evidence for a liquidity mechanism, and so our experiment does not support that explanation for the observed index inclusion premium. According to Pagano (1989), a negative relationship between turnover and spreads can indicate a more liquid market, which leads to higher asset prices. To test that idea, we compare the turnover of the included asset C between treatment *A2C* and *ABC*. Consistent with our previous analysis, we use each session as an observation but divide the turnover by the number of subjects in each session to control for the supply of assets available for trading. We do not find a significant difference in turnover for the included asset (p-value of 0.745 using the FP non-parametric test). Further, we find that the spread for the included asset does not decrease; indeed, the substantial increase in the spread shown in Table 5 (46.03 in *A2C* versus 17.54 in *ABC*) is statistically significant (FP p-value of 0.034).

For asset A and the excluded asset B, we do not find a significant change in spreads across treatments *ABC* and *A2C* (FP p-value of 0.123 and 0.183). For asset turnover, there is no significant change for asset A (p-value of 0.119), and a weak decrease for asset B (p-value of 0.06).

5 Conclusion

A leading explanation for the index inclusion premium is that demand for the constituent assets is downward sloping, as proposed by Shleifer (1986) and documented in the empirical literature, e.g., by Greenwood (2005) and Chang et al. (2015). Our laboratory experiment supports and further refines that explanation, while clearly ruling out other possible explanations such as signalling or increases in the liquidity of assets included in an index.

First, we report striking evidence for an index inclusion premium and a clear violation of the law of one price. Relative to our baseline *ABC* treatment, where the ETF index includes a unit of each of the three assets, when we replace asset B with one unit of an identical asset C in treatment $A2C$, we find that the price of the included asset C is significantly greater than in the baseline ABC treatment. Consistent with some archival studies, we also find a negative price impact on the excluded asset B.[11] We further show that in the $A2C$ treatment there is a positive order imbalance for asset C relative to asset B. Thus, our results confirm related empirical work showing that demand shocks to ETFs may be passed through to the constituent assets (Ben-David et al., 2018 and demonstrated by Box et al., 2021 for large ETF funds).

Second, we show that the index inclusion premium is not eliminated by allowing short-selling as in our treatment $A2C_{short}$, although short-selling does reduce order imbalances, prices, and spreads. Instead, we find that the spreads in treatment $A2C_{short}$ remain large enough so that the arbitrage trade of selling asset C and buying asset B remains unprofitable.

Third, we show that the mechanism driving the index inclusion premium is the strong demand for the ETF asset by traders seeking to diversify their initial portfolios in the direction of the market portfolio. Holding the ETF asset is a convenient, and perhaps cognitively less taxing, tool for that purpose. We find that there is a strong positive order imbalance for the ETF asset, and that most traders who buy the ETF asset employ a buy-and-hold strategy for the duration of the market. Indeed, we do not find any redemptions of the ETF asset, though there are trades of the ETF asset in the secondary market. The larger number of issues compared with redemptions observed in our laboratory study is consistent with ICI (2021) who report that the net share issuance of ETFs has increased steadily over the past 10 years in the United States. At some point investors' demand for ETF shares may reach some equilibrium fraction of the overall market, and so we might expect there to eventually be some smaller pressure on the prices of the constituent assets.

As noted, our laboratory experiment allows us to rule out other possible explanations for the index inclusion premium. The signalling hypothesis, that index inclusion serves as a signal of a stock's quality and growth potential and hence justifies higher prices, is ruled out by our experimental design, as there are no growth prospects for any of the constituent assets, included or excluded. We also do not find support for the liquidity hypothesis, that index inclusion leads to increased liquidity for a stock and narrower bid-ask spreads. Instead we find that turnover for the included asset C in treatment $A2C$ is no greater than in treatment ABC, while the spread for asset C

[11] In the field, it is difficult to find an excluded asset that is identical in state-contingent payoffs to the included asset (see the discussion in Wurgler and Zhuravskaya, 2002 for the S&P 500).

is actually greater in treatment *A2C* as compared with treatment *ABC*.

A novel feature of our experiment compared with other studies is that ETF shares appear endogenously without the presence of institutional investors. Additionally, we are able to study the effect of index-reconstitution for the entire market since the ETF share is offered in zero net supply without altering the aggregate risk in the economy. We are also able to consider the impact of index reconstitution on assets that were always included in the index, such as asset A. While ETFs can help traders to achieve the market portfolio, we show that the ETF index composition has important consequences for prices and market efficiency.

There remains much room for further work in experimental finance to provide more light on the mechanisms affecting asset prices, included or excluded from market indexes, complementing what can be inferred from archival data. For example, one might analyze the role of different investor types (active versus passive, or informed versus uninformed traders) or different ETF products that do not cover the entire market or that use leverage. We leave such topics for future research.

References

Amihud, Yakov and Haim Mendelson, "Liquidity, maturity, and the yields on US Treasury securities," *The Journal of Finance*, 1991, *46* (4), 1411–1425.

Asparouhova, Elena, Peter Bossaerts, Jernej Čopič, Brad Cornell, Jakša Cvitanić, and Debrah Meloso, "Competition in portfolio management: theory and experiment," *Management Science*, 2015, *61* (8), 1868–1888.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, "Do ETFs increase volatility?," *The Journal of Finance*, 2018, *73* (6), 2471–2535.

Bossaerts, Peter and Charles Plott, "Basic principles of asset pricing theory: Evidence from large-scale experimental financial markets," *Review of Finance*, 2004, *8* (2), 135–169.

Box, Travis, Ryan Davis, Richard Evans, and Andrew Lynch, "Intraday arbitrage between ETFs and their underlying portfolios," *Journal of Financial Economics*, 2021.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich, "Regression discontinuity and the price effects of stock market indexing," *The Review of Financial Studies*, 2015, *28* (1), 212–246.

Charness, Gary and Tibor Neugebauer, "A Test of the Modigliani-Miller Invariance Theorem and Arbitrage in Experimental Asset Markets," *The Journal of Finance*, 2019, *74* (1), 493–529.

Chen, Daniel L, Martin Schonger, and Chris Wickens, "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

Childs, Jason and Stuart Mestelman, "Rate-of-return parity in experimental asset markets," *Review of International Economics*, 2006, *14* (3), 331–347.

Chordia, Tarun and Avanidhar Subrahmanyam, "Order imbalance and individual stock returns: Theory and evidence," *Journal of Financial Economics*, 2004, *72* (3), 485–518.

Denis, Diane K, John J McConnell, Alexei V Ovtchinnikov, and Yun Yu, "S&P 500 index additions and earnings expectations," *The Journal of Finance*, 2003, *58* (5), 1821–1840.

Dhillon, Upinder and Herb Johnson, "Changes in the Standard and Poor's 500 List," *Journal of Business*, 1991, pp. 75–85.

Duffy, John, Jean Paul Rabanal, and Olga A Rud, "The impact of ETFs in secondary asset markets: Experimental evidence," *Journal of Economic Behavior & Organization*, 2021, *188*, 674–696.

_ , _ , and Olga Rud, "Market Experiments with Multiple Assets: A Survey," in Sascha Füllbrunn and Ernan Haruvy, eds., *Handbook of Experimental Finance*, Edward Elgar Publishing, 2022, pp. 213–224.

Fisher, Eric O'N and Frank S Kelly, "Experimental foreign exchange markets," *Pacific Economic Review*, 2000, *5* (3), 365–387.

Fisher, Ronald A., *The design of experiments*, Oliver and Boyd Ltd, 1935.

Foucault, Thierry, Marco Pagano, Ailsa Roell, and Ailsa Röell, *Market liquidity: theory, evidence, and policy*, Oxford University Press, 2013.

Greenwood, Robin, "Short-and long-term demand curves for stocks: theory and evidence on the dynamics of arbitrage," *Journal of Financial Economics*, 2005, *75* (3), 607–649.

Haruvy, Ernan and Charles Noussair, " The Effect of Short Selling on Bubbles and Crashes in Experimental Spot Asset Markets," *The Journal of Finance*, 2006, *61* (3), 1119–1157.

Hirshleifer, Jack and John G Riley, *The analytics of uncertainty and information*, Cambridge University Press, 1992.

ICI, *Fact Book*, Investment Company Institute, 2021.

Jain, Prem C, "The effect on stock price of inclusion in or exclusion from the S&P 500," *Financial Analysts Journal*, 1987, *43* (1), 58–65.

Kaul, Aditya, Vikas Mehrotra, and Randall Morck, "Demand curves for stocks do slope down: New evidence from an index weights adjustment," *The Journal of Finance*, 2000, *55* (2), 893–912.

Pagano, Marco, " Trading Volume and Asset Liquidity," *The Quarterly Journal of Economics*, 1989, *104* (2), 255–274.

Pitman, Edwin James George, "Significance tests which may be applied to samples from any populations: III. The analysis of variance test," *Biometrika*, 1938, *29* (3/4), 322–335.

Sharpe, William F, "Capital asset prices: A theory of market equilibrium under conditions of risk," *The Journal of Finance*, 1964, *19* (3), 425–442.

Shleifer, Andrei, "Do demand curves for stocks slope down?," *The Journal of Finance*, 1986, *41* (3), 579–590.

Wurgler, Jeffrey and Ekaterina Zhuravskaya, "Does arbitrage flatten demand curves for stocks?," *The Journal of Business*, 2002, *75* (4), 583–608.

Appendix for online publication only

A Derivation of CAPM prices

Assets $\theta = \{A, B, C\}$ yield a state-dependent terminal dividend payout represented by matrix D, where the three states, X, Y and Z are equiprobable. The dividends have a mean μ and covariance Δ.

$$D = \begin{bmatrix} & X & Y & Z \\ A & 0 & 120 & 120 \\ B & 90 & 0 & 90 \\ C & 90 & 0 & 90 \end{bmatrix}, \mu = \begin{bmatrix} 80 \\ 60 \\ 60 \end{bmatrix}$$

Deviations from the mean are:

$$D - \mu = \begin{bmatrix} & X & Y & Z \\ A & -80 & 40 & 40 \\ B & 30 & -60 & 30 \\ C & 30 & -60 & 30 \end{bmatrix}$$

So,

$$\Delta = \begin{bmatrix} 3200 & -1200 & -1200 \\ -1200 & 1800 & 1800 \\ -1200 & 1800 & 1800 \end{bmatrix}$$

Let z denote the (post-trade) net demand. To determine the equilibrium price p of asset θ, we assume mean-variance utility $U = E(W) - \frac{b}{2}Var(W)$, where W is wealth

$$W = D \cdot z - p \cdot z \tag{A.1}$$

The market clearing condition implies that the total net demand of assets is equal to the fixed supply of assets in the economy, Q,

$$\sum z = Q \tag{A.2}$$

Thus, plugging W in the utility function and taking the FOC w.r.t z, we obtain

$$\mu - p - b\Delta z = 0 \tag{A.3}$$

Adding up the individual demands, we find

$$\sum \mu - \sum p - b\Delta \sum z = 0 \tag{A.4}$$

If we define $\bar{Q}$ as the total supply of assets per capita, then equilibrium prices are

$$p^*_\theta = \mu - b\Delta\bar{Q}, \tag{A.5}$$

Assuming $b = 0.001$, and $\bar{Q} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$, we obtain $p = (78, 55, 55)$.

Note that equation (A.5) is the same as equation (35) in Greenwood (2005).

B Experimental Instructions

Treatment *ABC*

Welcome to this experiment in market decision-making. Each participant is guaranteed $7 for showing up and completing today's session. In addition, you can earn points based on the decisions that you make which will be converted into additional dollar earnings. Your total earnings will be paid to you in US dollars at the end of today's session.

Kindly silence all electronic devices and do not talk with other participants for the duration of today's session. If you have any questions, or need assistance of any kind, please raise your hand.

General Information

This experiment consists of 7 separate rounds. Each round lasts 5 minutes. In each round, participants can trade (buy and/or sell) four assets A, B, C and D in exchange for experimental cash which will be converted into points (at a rate explained below). Following completion of the session, the computer program will randomly select one of the seven rounds to compute your final payment. The total points you earn in the selected round will be converted into your real money earnings at the rate of 1 point=

$0.04.

Prior to playing for points, you will have an additional 3 practice rounds, where each round is 3 minutes. These 3 rounds will not count toward your final earnings.

Market Description

Each market consists of 10-16 participants, who can trade up to four types of assets: A, B, C and D. You will enter each market with a loan of 210 in experimental cash which you will have to pay back at the end of the round. You will also begin each market with some units of asset A, asset B, and asset C, as will be clearly revealed on your computer screen. In every round you can buy or sell (trade) assets for experimental cash (henceforth, "cash").

There are four markets in each round, one for each of the four assets. You can be a buyer or a seller or both a buyer and a seller (a trader) of each asset in each of the four markets. To buy a unit, you submit a bid (buying price) and to sell a unit, you submit an ask (selling price). A bid or an ask can be any positive number up to some constraints. You can only buy or sell one unit at a time. Please note that your asset inventories cannot drop below zero. In other words, you cannot sell more assets than you have. Also, your bid times the number of units you want to buy of any asset cannot exceed your cash balance.

Asset D is a composite asset which is formed as a combination of 1 unit of asset A and 1 unit of asset B and 1 unit of asset C. At the start of each round, you will not hold any D assets, but a computerized trader may offer units of asset D for sale following the start of trade in each round. You can bid for D at any point in the market, and you can sell D, if you have it in your portfolio.

The computerized trader will offer the composite asset D for sale (request it for purchase) only when underlying asks (bids) for the composite asset D exist. For example, if the computerized trader sees asks for asset A only (that is, no one is offering to sell asset B or C), then it cannot offer asset D for sale. However, if all three assets are available for sale in the market, in sufficient quantities (according to the composition) then the computerized trader will offer D for sale at the following ask:

$$Ask(D) = Ask(A) + Ask(B) + Ask(C)$$

where Ask(A), Ask(B) and Ask(C) are currently the best available asks. Therefore, a computerized trader will offer asset D for sale only when the asks for underlying assets exist. Similarly, the computerized trader will submit a bid for D, when the underlying asset bids exist (at the current best available bids).

A transaction in any market occurs when the highest bid crosses the lowest ask– that is, when there is a price that someone is willing to pay which is at least as high as the lowest ask, or when the price someone is willing to accept is lower than the highest bid. Bids are sorted in the book from highest to lowest, while asks are sorted from lowest to highest. You can also transact by double clicking on the bid or ask that appears in the order book on your screen, and confirming the transaction.

At the end of each trading round, each asset yields a payoff according to the table shown in Figure 1, which will also appear in the bottom right of your trading screen, see Figure 2. The payoff for each asset depends on which of three possible "states" X, Y, or Z occurs. Each possible state (X, Y and Z) occurs with an equal 1/3chance. The computer will randomly select a state and the selected state will be shown to you at the end of each round. The state selected will be randomly and independently chosen for each round. Therefore, the draw of a state in one round does not have any effect on the state drawn in subsequent rounds.

State	X	Y	Z
Probability	1/3	1/3	1/3
A payoff	0	120	120
B payoff	90	0	90
C payoff	90	0	90
D payoff	180	120	300

Figure 1: Payoff states

Your total points in a round (earnings) is determined by the following formula:

Total points = The value of your total assets held in the chosen round + Final cash balance - Loan of 210:

where:

Value of your total assets held in the chosen round= Units of A * A Payoff + Units of B * B Payoff + Units of C * C Payoff + Units of D * D Payoff.

(Figure 2 in instructions: similar to Figure 1 in the main paper but with no orders submitted).

Quiz (coded in oTree)

1. What is the value of asset C in state Z (please refer to the state payoff table above)?

    (a) 120

    (b) 0

    (c) 90

    (d) 180

2. What is the value of asset D in state X?

    (a) 180

    (b) 90

    (c) 300

    (d) 0

3. What is the value of asset A in state Y?

    (a) 180

    (b) 120

    (c) 300

    (d) 0

4. If you hold 1 D asset and 2 B assets, and asset D = 1A + 1B + 1C and state Y is realized, what is the value of your portfolio?

    (a) 0

    (b) 120

    (c) 300

    (d) 90

5. If in one round state X was drawn, then what is the probability that Y occurs in another round?

    (a) The draw of states is independent in each round, therefore the probability of observing Y in any round is 1/3.

(b) The draw of states is not independent of each other, therefore the probability of observing Y in any round is 2/3

(c) It cannot be determined.

6. For a computerized trader to offer asset D for sale, which is composed of 1A + 1B +1C, which asks need to exist in the market?

   (a) Asks for asset C

   (b) Asks for asset B

   (c) Asks for asset A

   (d) Asks for A, B, and C

Treatment *A2C*

Welcome to this experiment in market decision-making. Each participant is guaranteed $7 for showing up and completing today's session. In addition, you can earn points based on the decisions that you make which will be converted into additional dollar earnings. Your total earnings will be paid to you in US dollars at the end of today's session.

Kindly silence all electronic devices and do not talk with other participants for the duration of today's session. If you have any questions, or need assistance of any kind, please raise your hand.

General Information

This experiment consists of 7 separate rounds. Each round lasts 5 minutes. In each round, participants can trade (buy and/or sell) four assets A, B, C and D in exchange for experimental cash which will be converted into points (at a rate explained below). Following completion of the session, the computer program will randomly select one of the seven rounds to compute your final payment. The total points you earn in the selected round will be converted into your real money earnings at the rate of 1 point= $0.04.

Prior to playing for points, you will have an additional 3 practice rounds, where each round is 3 minutes. These 3 rounds will not count toward your final earnings.

Market Description

Each market consists of 10-16 participants, who can trade up to four types of assets: A, B, C and D. You will enter each market with a loan of 210 in experimental cash which you will have to pay back at the end of the round. You will also begin each market with some units of asset A, asset B, and asset C, as will be clearly revealed on your computer screen. In every round you can buy or sell (trade) assets for experimental cash (henceforth, "cash").

There are four markets in each round, one for each of the four assets. You can be a buyer or a seller or both a buyer and a seller (a trader) of each asset in each of the four markets. To buy a unit, you submit a bid (buying price) and to sell a unit, you submit an ask (selling price). A bid or an ask can be any positive number up to some constraints. You can only buy or sell one unit at a time. Please note that your asset inventories cannot drop below zero. In other words, you cannot sell more assets than you have. Also, your bid times the number of units you want to buy of any asset cannot exceed your cash balance.

Asset D is a composite asset which is formed as a combination of 1 unit of asset A and 2 units of asset C. At the start of each round, you will not hold any D assets, but a computerized trader may offer units of asset D for sale following the start of trade in each round. You can bid for D at any point in the market, and you can sell D, if you have it in your portfolio.

The computerized trader will offer the composite asset D for sale (request it for purchase) only when underlying asks (bids) for the composite asset D exist. For example, if the computerized trader observes asks for asset A only (that is, no one is offering asset C for sale), then it cannot offer asset D for sale. However, if both assets are available in the market, in sufficient quantities (according to the composition) then the computerized trader will offer D for sale at the following ask:

$$Ask(D) = Ask(A) + Ask(C) + Ask(C)$$

where Ask(A), Ask(C) and Ask(C) are currently the best available asks for 1 unit of Asset A and 2 units of Asset C. Therefore, a computerized trader will offer asset D for sale only when the asks for underlying assets exist. Similarly, the computerized trader will submit a bid for D, when the underlying asset bids exist.

A transaction in any market occurs when the highest bid crosses the lowest ask– that is, when there is a price that someone is willing to pay which is at least as high as the lowest ask, or when the price someone is willing to accept is lower than the highest bid. Bids are sorted in the book from highest to lowest, while asks are sorted from lowest to highest. You can also transact by double clicking on the bid or ask that appears in the order book on your screen, and confirming the transaction.

At the end of each trading round, each asset yields a payoff according to the table shown in Figure 1, which will also appear in the bottom right of your trading screen, see Figure 2. The payoff for each asset depends on which of three possible "states" X, Y, or Z occurs. Each possible state (X, Y and Z) occurs with an equal 1/3 chance. The computer will randomly select a state and the selected state will be shown to you at the end of each round. The state selected will be randomly and independently chosen for each round. Therefore, the draw of a state in one round does not have any effect on the state drawn in subsequent rounds.

State	X	Y	Z
Probability	1/3	1/3	1/3
A payoff	0	120	120
B payoff	90	0	90
C payoff	90	0	90
D payoff	180	120	300

Figure 1: Payoff states

Your total points in a round (earnings) is determined by the following formula:

Total points =

The value of your total assets held in the chosen round + Final cash balance - Loan of 210:

where:

Value of your total assets held in the chosen round= Units of A * A Payoff + Units of B * B Payoff + Units of C * C Payoff + Units of D * D Payoff.

(Figure 2 in instructions: similar to Figure 1 in the main paper but with no orders submitted).

Quiz (coded in oTree)

1. What is the value of asset C in state Z (please refer to the state payoff table above)?

   (a) 120

   (b) 0

(c) 90

(d) 180

2. What is the value of asset D in state X?

   (a) 180

   (b) 90

   (c) 300

   (d) 0

3. What is the value of asset A in state Y?

   (a) 180

   (b) 120

   (c) 300

   (d) 0

4. If you hold 1 D asset and 2 B assets, and asset D = 1A + 2C and state Y is realized, what is the value of your portfolio?

   (a) 0

   (b) 120

   (c) 300

   (d) 90

5. If in one round state X was drawn, then what is the probability that Y occurs in another round?

   (a) The draw of states is independent in each round, therefore the probability of observing Y in any round is 1/3.

   (b) The draw of states is not independent of each other, therefore the probability of observing Y in any round is 2/3

   (c) It cannot be determined.

6. For a computerized trader to offer asset D for sale, which is composed of 1A + 2C, which asks need to exist in the market?

(a) Asks for asset C

(b) Asks for 1 A and 1C

(c) Asks for 1A and 2C

Treatment $\boldsymbol{A2C}_{short}$

Welcome to this experiment in market decision-making. Each participant is guaranteed $7 for showing up and completing today's session. In addition, you can earn points based on the decisions that you make which will be converted into additional dollar earnings. Your total earnings will be paid to you in US dollars at the end of today's session.

Kindly silence all electronic devices and do not talk with other participants for the duration of today's session. If you have any questions, or need assistance of any kind, please raise your hand.

General Information

This experiment consists of 7 separate rounds. Each round lasts 5 minutes. In each round, participants can trade (buy and/or sell) four assets A, B, C and D in exchange for experimental cash which will be converted into points (at a rate explained below). Following completion of the session, the computer program will randomly select one of the seven rounds to compute your final payment. The total points you earn in the selected round will be converted into your real money earnings at the rate of 1 point= 0.04.

Prior to playing for points, you will have an additional 3 practice rounds, where each round is 3 minutes. These 3 rounds will not count toward your final earnings.

Market Description

Each market consists of 10-16 participants, who can trade up to four types of assets: A, B, C and D. You will enter each market with a loan of 210 in experimental cash which you will have to pay back at the end of the round. You will also begin each market with some units of asset A, asset B, and asset C, as will be clearly revealed on your computer screen. In every round you can buy or sell (trade) assets for experimental cash (henceforth, "cash").

There are four markets in each round, one for each of the four assets. You can be a buyer or a seller or both a buyer and a seller (a trader) of each asset in each of the four markets. To buy a unit, you submit a bid (buying price) and to sell a unit, you submit an ask (selling price). A bid or an ask can be any positive number up to two decimal places. You can only buy or sell one unit at a time. Also, your bid times the number of units you want to buy of any asset cannot exceed your cash balance.

You can sell an asset that you do not have (this is allowed only for assets A, B and C). In this case your asset holdings for this asset will show as being negative. That is, if you sell 1 unit of asset B, which you do not have, then your asset holdings for asset B will show up as -1 (or -2, if you sell 2 units of asset B that you do not have).

If you do not balance your portfolio by the end of the round so that you are no longer negative in any assets A, B or C, then you will be charged the payoff value of any asset(s) for which you have a negative balance. The payoff value depends on the states in Figure 1 (as explained below) and will be deducted from your portfolio value at the end of the round. For example, if your position for asset A is -1 at the end of the round, then we will subtract the actual payoff value of asset A from your earnings for that round.

Asset D is a composite asset which is formed as a combination of 1 unit of asset A and 2 units of asset C. At the start of each round, you will not hold any D assets, but a computerized trader may offer units of asset D for sale following the start of trade in each round. You can bid for D at any point in the market, and you can sell D, if you have it in your portfolio.

The computerized trader will offer the composite asset D for sale (request it for purchase) only when underlying asks (bids) for the composite asset D exist. For example, if the computerized trader observes asks for asset A only (that is, no one is offering asset C for sale), then it cannot offer asset D for sale. However, if both assets are available in the market, in sufficient quantities (according to the composition) then the computerized trader will offer D for sale at the following ask:

$$Ask(D) = Ask(A) + Ask(C) + Ask(C)$$

where Ask(A), Ask(C) and Ask(C) are currently the best available asks for 1 unit of Asset A and 2 units of Asset C. Therefore, a computerized trader will offer asset D for sale only when the asks for underlying assets exist. Similarly, the computerized trader will submit a bid for D, when the underlying asset bids exist.

A transaction in any market occurs when the highest bid crosses the lowest ask– that is, when there is a price that someone is willing to pay which is at least as high as the lowest ask, or when the price someone is willing to accept is lower than the highest bid. Bids are sorted in the book from highest to lowest, while asks are sorted from lowest to highest. You can also transact by double clicking on the bid or ask that appears in the order book on your screen, and confirming the transaction.

At the end of each trading round, each asset yields a payoff according to the table shown in Figure 1, which will also appear in the bottom right of your trading screen, see Figure 2. The payoff for each asset depends on which of three possible "states" X, Y, or Z occurs. Each possible state (X, Y and Z) occurs with an equal 1/3 chance. The computer will randomly select a state and the selected state will be shown to you at the end of each round. The state selected will be randomly and independently chosen for each round. Therefore, the draw of a state in one round does not have any effect on the state drawn in subsequent rounds.

State	X	Y	Z
Probability	1/3	1/3	1/3
A payoff	0	120	120
B payoff	90	0	90
C payoff	90	0	90
D payoff	180	120	300

Figure 1: Payoff states

Your total points in a round (earnings) is determined by the following formula:

Total points =

The value of your total assets held in the chosen round + Final cash balance - Loan of 210:

where:

Value of your total assets held in the chosen round= Units of A * A Payoff + Units of B * B Payoff + Units of C * C Payoff + Units of D * D Payoff.

(Figure 2 in instructions: similar to Figure 1 in the main paper but with no orders submitted).

Quiz (coded in oTree)

The quiz is similar to *A2C* plus the following question:

1. If you finish the round with -2 of asset B and state X is drawn, this means that:

(a) You bought 2 more units of B than you had in your portfolio and 2*90 will be credited to your earnings

(b) You sold 2 more units of B than you had in your portfolio and 0 will be deducted/credited to your earnings

(c) You sold 2 more units of B than you had in your portfolio and 2*90 will be deducted from your earnings

C Summary of Session Data

Sessions for treatment ABC

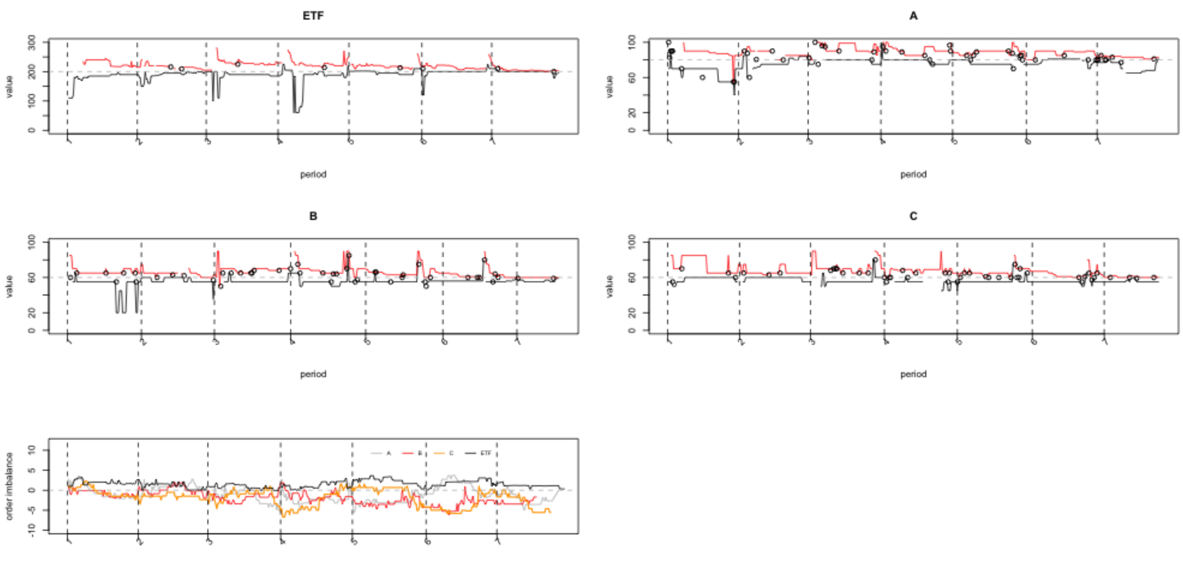

Figure C.1: Best bids/offers, prices and order imbalance for ABC treatment (session 1)

Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

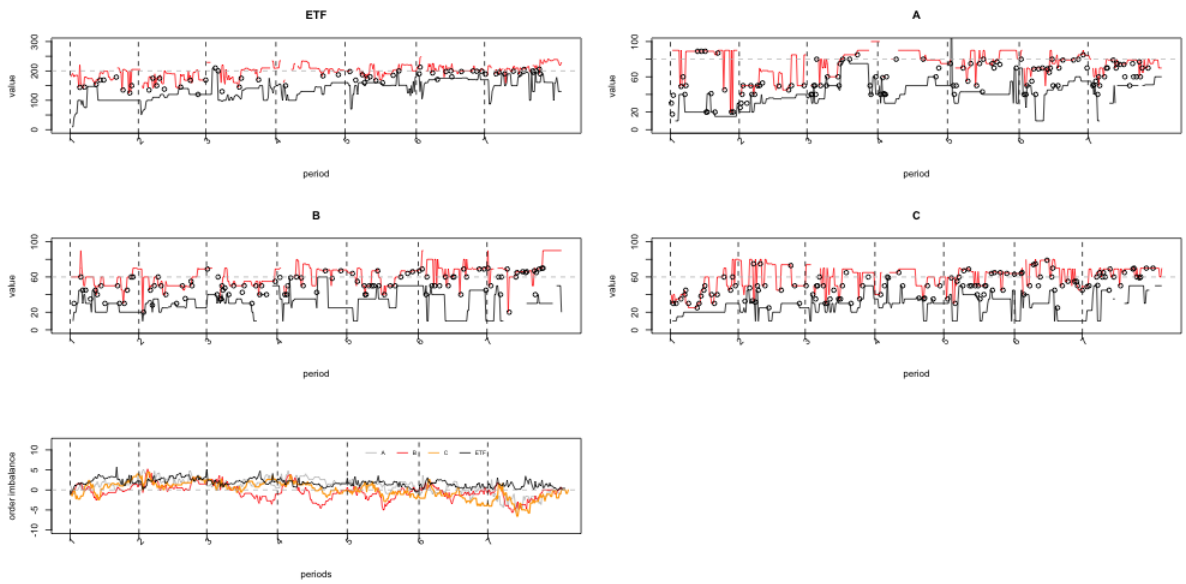

Figure C.2: Best bids/offers, prices and order imbalance for ABC treatment (session 2)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

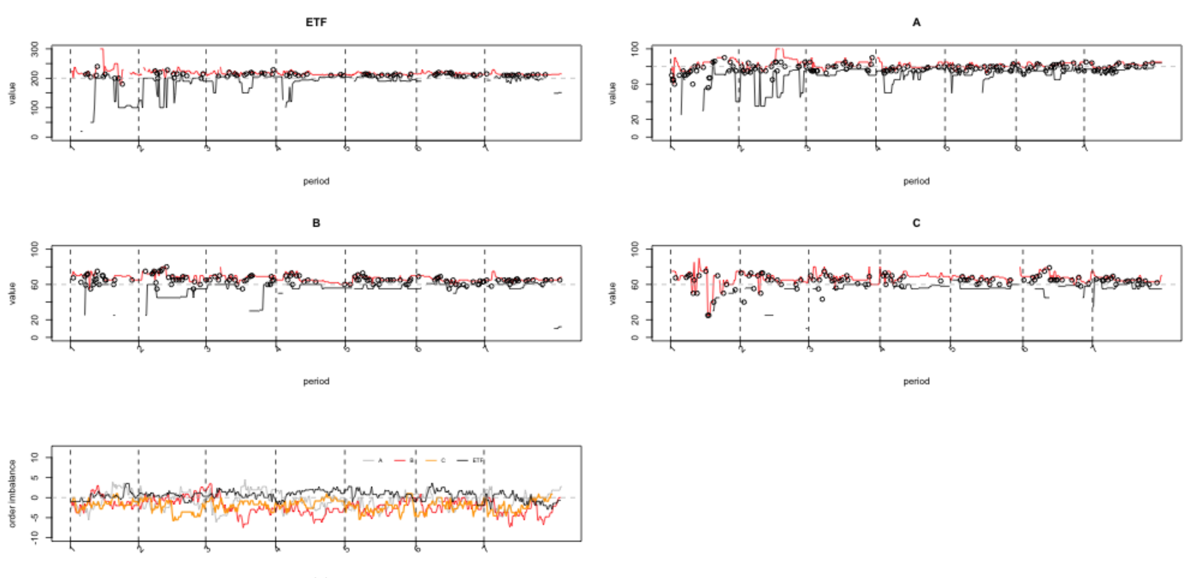

Figure C.3: Best bids/offers, prices and order imbalance for ABC treatment (session 3)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

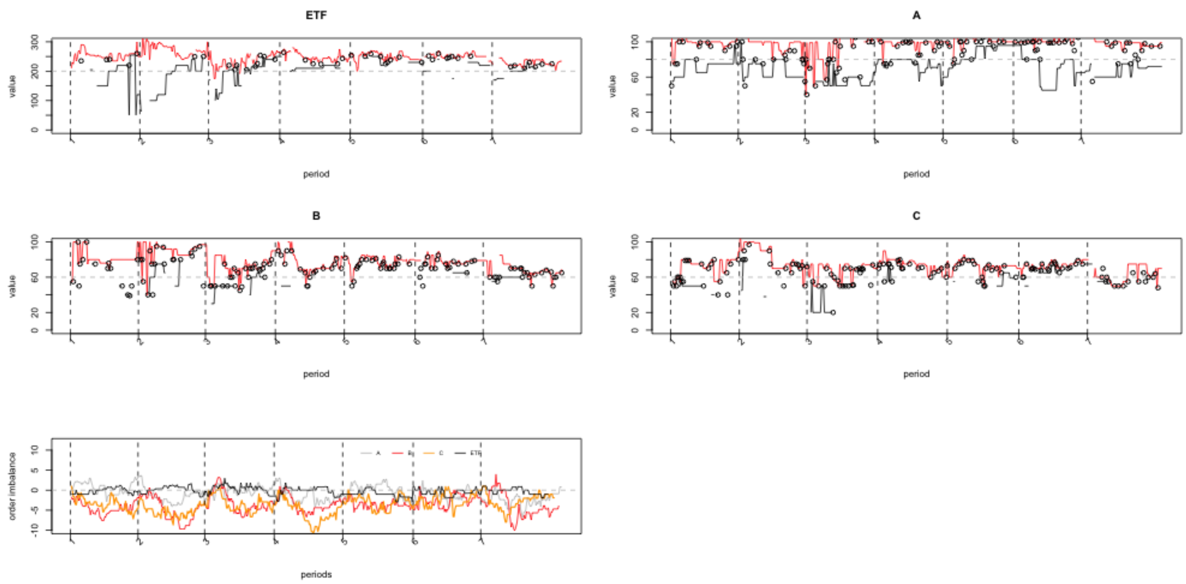

Figure C.4: Best bids/offers, prices and order imbalance for ABC treatment (session 4)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

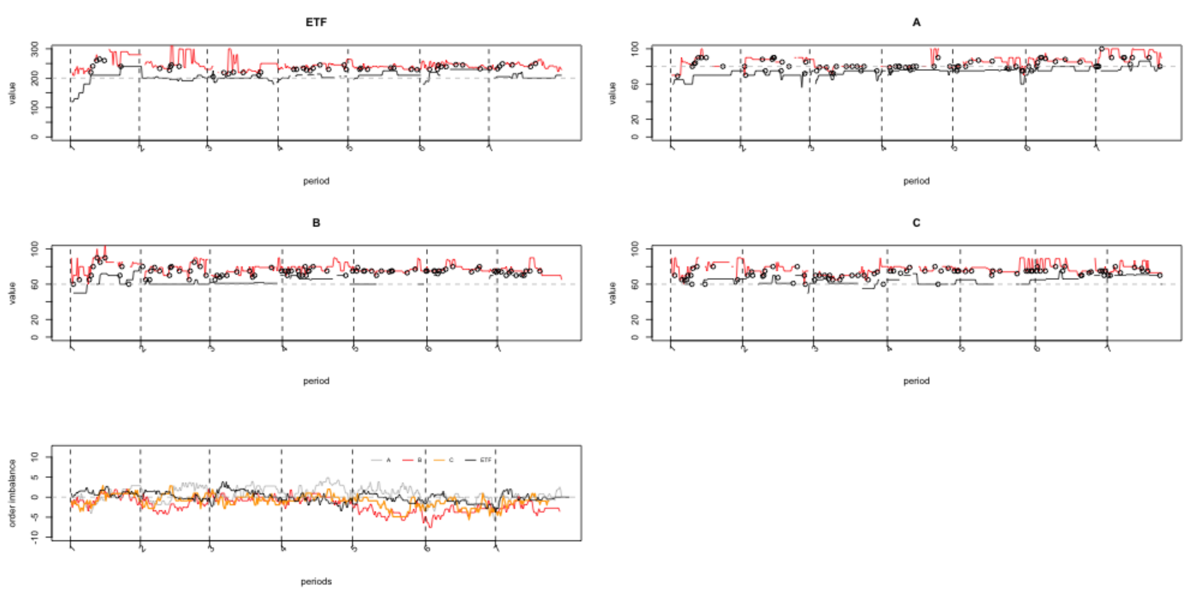

Figure C.5: Best bids/offers, prices and order imbalance for ABC treatment (session 5)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

Sessions for treatment *A2C*

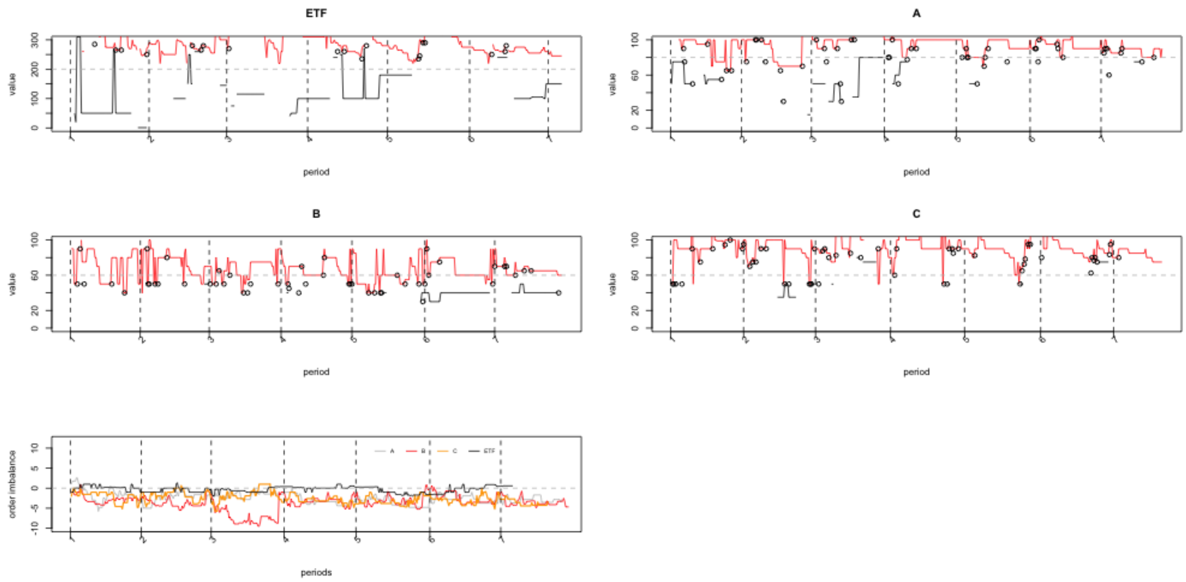

Figure C.6: Best bids/offers, prices and order imbalance for *A2C* treatment (session 1)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: *ETF*, *A*, *B* and *C*. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets *A* (gray), asset *B* (red), asset *C* (orange), and the *ETF* (black).

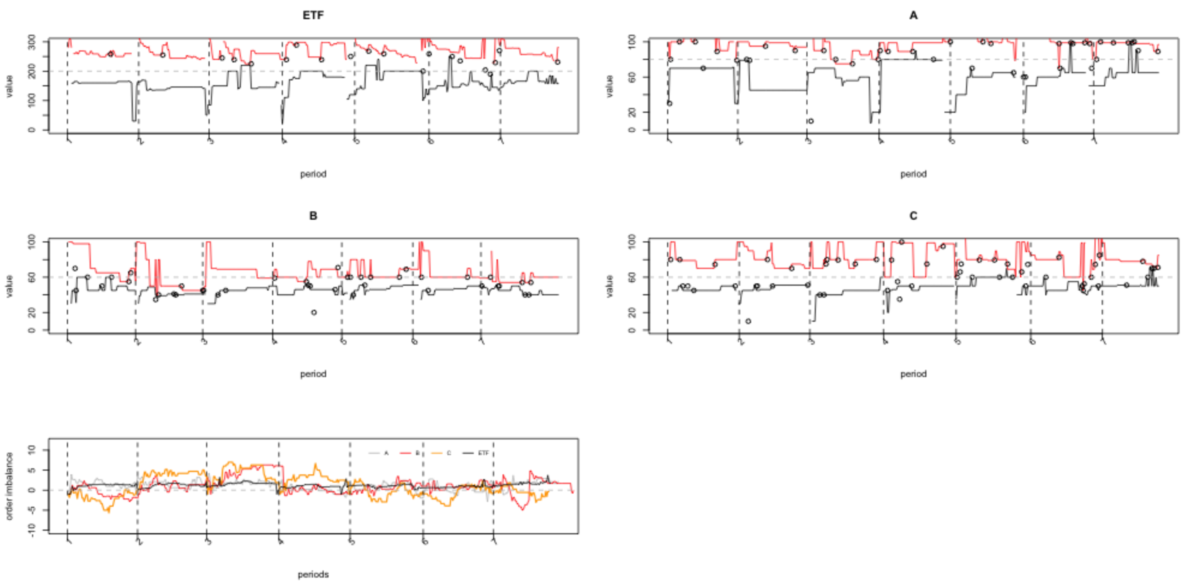

Figure C.7: Best bids/offers, prices and order imbalance for *A2C* treatment (session 2)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

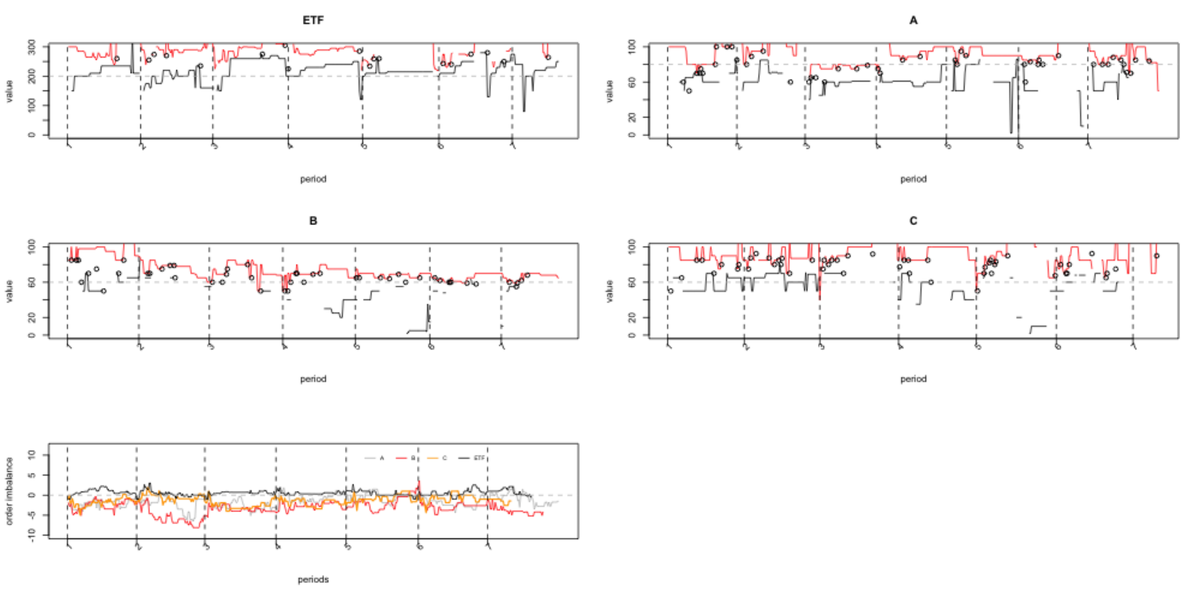

Figure C.8: Best bids/offers, prices and order imbalance for *A2C* treatment (session 3)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

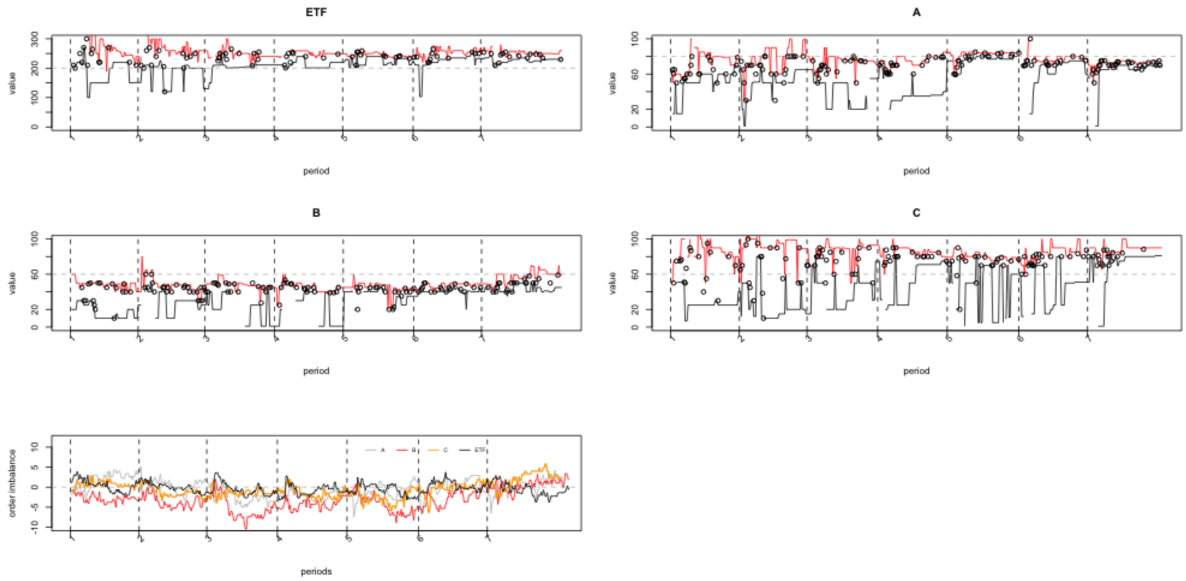

Figure C.9: Best bids/offers, prices and order imbalance for *A2C* treatment (session 4)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

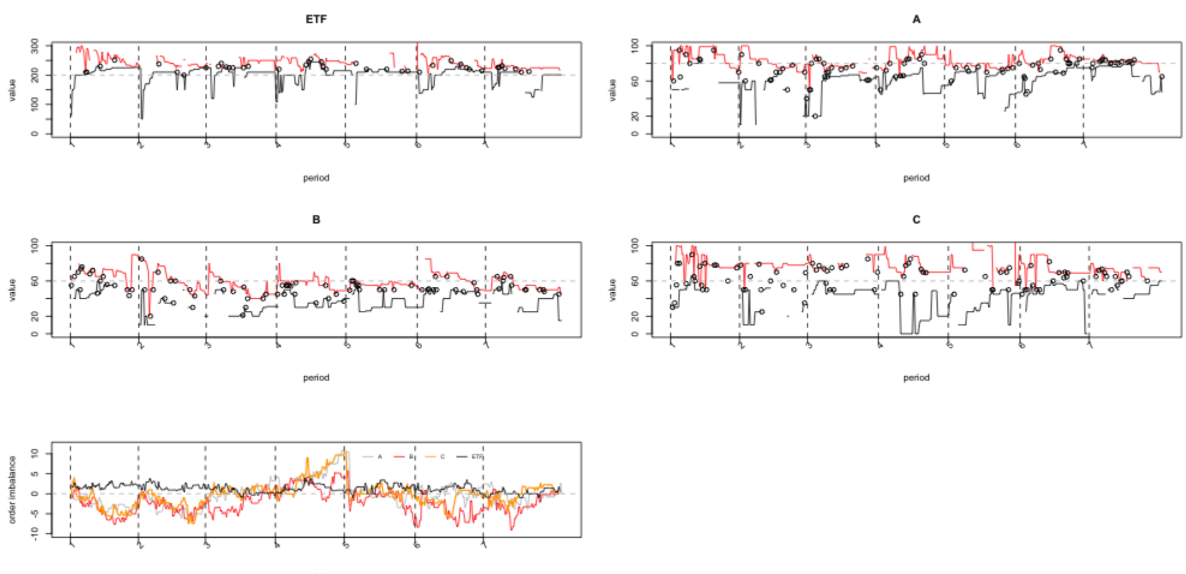

Figure C.10: Best bids/offers, prices and order imbalance for *A2C* treatment (session 5)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

Sessions for treatment $A2C_{short}$

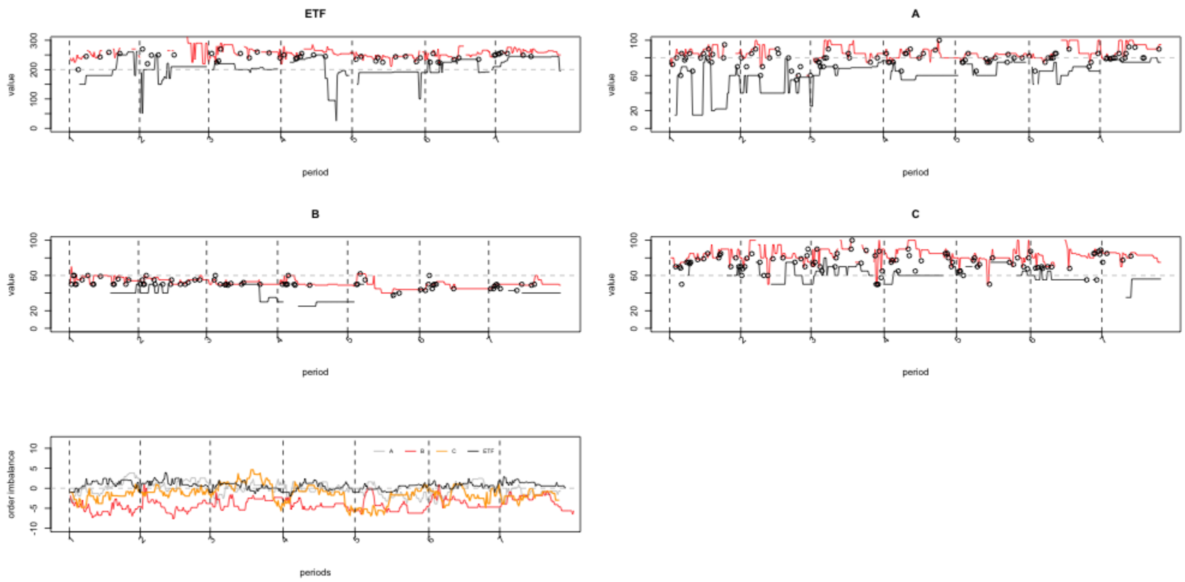

Figure C.11: Best bids/offers, prices and order imbalance for $A2C_{short}$ treatment (session 1)

Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

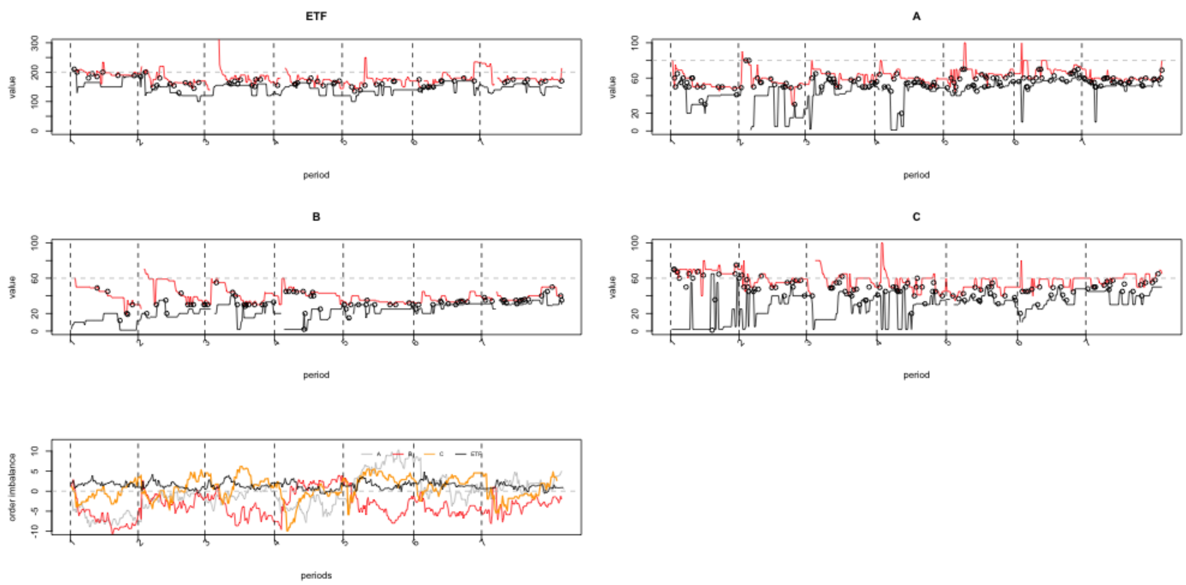

Figure C.12: Best bids/offers, prices and order imbalance for $A2C_{short}$ treatment (session 2)

Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

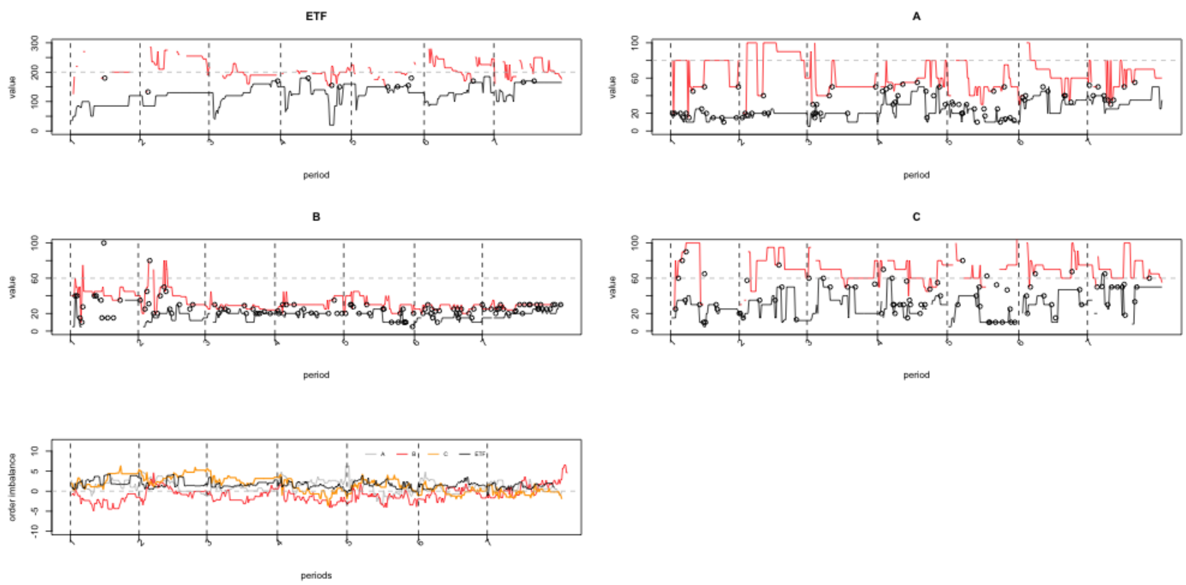

Figure C.13: Best bids/offers, prices and order imbalance for $A2C_{short}$ treatment (session 3)

Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

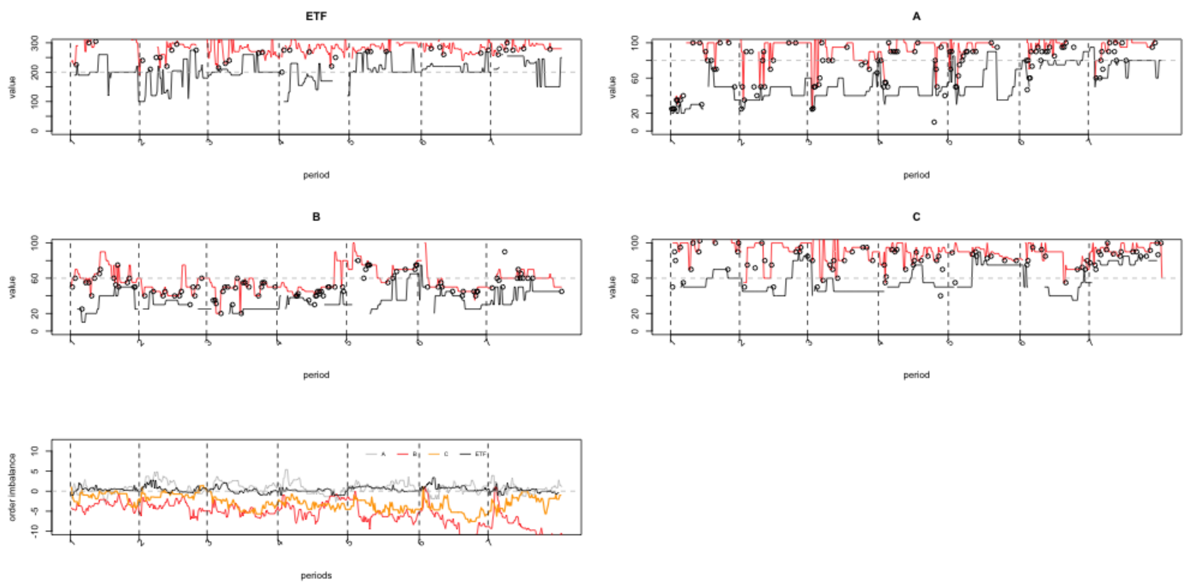

Figure C.14: Best bids/offers, prices and order imbalance for $A2C_{short}$ treatment (session 4)

Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

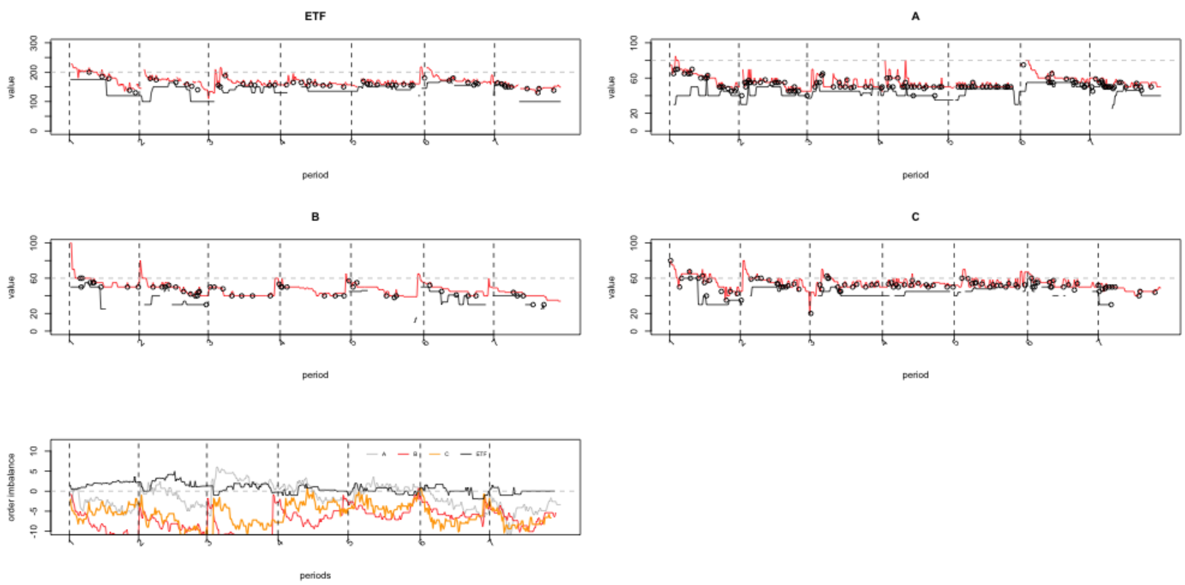

Figure C.15: Best bids/offers, prices and order imbalance for $A2C_{short}$ treatment (session 5)

Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF, A, B and C. Each tick corresponds to four seconds, and the dashed vertical lines denote the start of a trading period. The horizontal dashed line is the expected value of the asset. The bottom row presents the order imbalance specified in equation (3), where the four series correspond to order imbalances for assets A (gray), asset B (red), asset C (orange), and the ETF (black).

D Additional empirical results

Table D.1: ETF transactions

Session	1	2	3	4	5
			ABC		
Issues	9	57	41	74	53
Secondary	0	6	12	13	0
			A2C		
Issues	26	19	20	63	38
Secondary	0	1	2	32	1
			$A2C_{short}$		
Issues	44	54	12	35	55
Secondary	7	14	1	4	2

Note: We count the number of total ETF transactions in each session, and classify the transactions as new issues (AP bot), or secondary trades (human subjects). In our sample, the AP never redeemed ETFs.

Further results

Result 5: *Towards the end of each trading period in treatments* A2C *and* A2C_{short}*, the best ask prices for the included asset* C *are significantly higher than the best ask prices for the excluded asset* B*.*

When a new ETF share is issued, the AP-bot submits market orders for the underlying assets. Table D.2 suggests that, late in the trading period, bot orders for the included asset C are executed at an average price of around 89 in treatment *A2C* and at an average price of around 70 in treatment $A2C_{short}$, which is greater than its expected value of 60. The lower ask prices in treatment $A2C_{short}$ can be explained by the greater supply of asset C when short-selling is allowed. We also observe a larger divergence in the bid prices for assets B and C in treatment $A2C_{short}$. Overall, excluding asset B results in an increase of ask prices for asset C which is larger than the increase in bid prices; resulting in a higher spread for C compared to the baseline treatment *ABC* (see Table 5). For asset B, the excluded asset, we do not see major changes in the spread since both ask and bid prices drop by a similar amount.

Short-selling should help subjects exploit arbitrage opportunities between the prices of identical assets B and C. To investigate the profitability of selling C and buying B, the last row of Table D.2 reports the average difference between the best bid price

Table D.2: Best bid and asks (30 sec before market closed)

	Ask			Bid		
	ABC	*A2C*	*A2C*$_{short}$	*ABC*	*A2C*	*A2C*$_{short}$
A	87	95	72	66	58	49
B	73	61	43	50	38	30
C	65	89	70	55	46	45
ETF	224	281	257	191	183	165
Bid *C* - Ask *B*	-17	-11	-1	–	–	–

Note: We take a snapshot of the best bids and asks 30 seconds prior to the end of each trading period.

for asset *C* and the best ask price for asset *B*; a positive difference would indicate an arbitrage opportunity.

Result 6: *There are limited arbitrage opportunities: the average difference between the best bid for asset C and the best ask for asset B is negative or zero.*

We observe a negative difference between the best bid for *C* and the best ask for *B* in treatment *ABC* of about -17 (which is always negative in all sessions, *p*-value of 0.071). This difference becomes smaller in the other two treatments, driven by lower ask prices for the excluded asset *B*, and is close to zero in the *A2C*$_{short}$ treatment.

To investigate more directly whether subjects engage in arbitrage by selling asset *C* and buying asset *B* we count the number of subjects who bought one identical asset and sold the other, relative to the total number of subjects present in the market. These findings are summarized in Table D.3. Overall, we do not observe significant differences in the number of subjects selling asset *C* and buying asset *B* across the three treatments. This is consistent with Result 5, which suggests that there are few profitable arbitrage opportunities. We also look at the number of subjects who buy asset *C* and sell asset *B* to check the consistency of our results. In treatment *ABC*, we observe a similar fraction of such subjects (around 1/3) which confirms the fact that *B* and *C* are identical, so one should expect a similar fraction of subjects buying and selling the identical assets. For the other two treatments, we find some differences. We see that the fraction of subjects who buy *C* and sell *B* decreases from 0.34 in treatment *ABC* to 0.21 in treatment *A2C*, and to 0.24 in treatment *A2C*$_{short}$. This can be explained by the fact that the AP-bot demands more units of asset *C* to create an ETF in these two treatments, which decreases the number of units of asset *C* that are available in the market, and/or that it is not profitable to engage in buying asset C and selling asset B compared to buying asset B and selling asset C given the index premium documented earlier.

Table D.3: Subjects trades (fraction of total Ss per period)

	ABC	$A2C$	$A2C_{short}$
Buy B & Sell C	0.33	0.33	0.37
Buy C & Sell B	0.34	0.21	0.24

Note: We count the number of subjects who buy asset B and sell asset C (and buy C and sell B) per trading period and present the data in terms of the total number of subjects in the market.

Result 7: *The AP-bot trader generates a small profit from trading in the primary market.*

The AP-bot trader can earn profit via the ETF bid and ask spread. That is, the AP-bot can create an ETF share at a higher ask price, and redeem it at the lower bid price. In our sample, as documented above, we do not observe any ETF redemptions, and therefore the AP-bot does not generate profit by buying low and selling high. Alternatively, it could be the case that players bid for the ETF share at a higher value that the sum of the ask prices of the underlying assets. If the bid has a time-priority, then $P_{ETF} > NAV$ and the AP-bot makes a profit.

Table D.4: Average AP-bot trader profit and ETF units traded

Treatment	Profit (1)	ETF trades (2)	Profit per unit (1/2)	$P_{ETF} = NAV$ (%)
ABC	10.17	6.68	1.52	83.35
$A2C$	13.66	4.74	2.88	80.69
$A2C_{short}$	7.00	5.71	1.23	79.00

Note: (1) reports average AP-bot profit per period, (2) reports the per-period average number of ETF units traded when one side of the market is a bot. Profit per unit divides the column (1) entry by the column (2) entry. $P_{ETF} = NAV$ reports the fraction of all ETF trades where price is equal to the NAV.

Table D.4 presents the AP-bot profit and arbitrage activity. The first column of Table D.4 presents the average profit: it appears that the bot profit is highest in the $A2C$ treatment, and lowest in the $A2C_{short}$ treatment. On average, we find that there are between 5 or 7 ETF units traded per trading period with the greatest number of ETF trades occurring in the ABC treatment. Around 80 percent of those units (the last column) are traded without any arbitrage opportunities, that is, the ETF price is equal to the NAV ($P_{ETF} = NAV$) in the primary market. The AP-bot's profit per transaction is quite small, averaging about 1.2 per unit in treatment $A2C_{short}$, and increasing slightly to an average of 2.9 per unit in the treatment $A2C$.