

The impact of ETF index inclusion on stock prices

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

March 22, 2022

Abstract

A growing body of evidence suggests that assets included in market indexes such as the S&P 500 trade at a premium relative to other assets. In this paper we look for evidence of such an index inclusion premium in a carefully controlled laboratory experiment. Our environment involves three assets and an Exchange Traded Fund (ETF) index asset. We model Authorized Participants (APs) as bots that create and redeem ETF shares by scanning the order books of the underlying assets. In one treatment, all three assets are included in the ETF index asset. In a second treatment, one of the three assets is excluded from the ETF index and is replaced by a second unit of one of the included assets; the included and excluded assets have identical fundamental values enabling a clean test of whether or not there exists an index inclusion premium. We consider a further variant of the excluded asset treatment where short-selling is allowed. We find that: (i) inclusion of an asset in the ETF index results in a substantial index premium, (ii) this result is tied to an order imbalance which arises when an identical asset is excluded from the index, and (iii) the premium and order imbalance persist even if short-selling is allowed.

Keywords: index premium, ETFs, experimental finance

*John Duffy (duffy@uci.edu) is at University of California, Irvine; Dan Friedman (dan@ucsc.edu) is at University of California Santa Cruz, and University of Essex; Jean Paul Rabanal (jean-paul.rabanal@uis.no) and Olga Rud (olga.rud@gmail.com) are at the University of Stavanger. We gratefully acknowledge research support for this project from the IFREE Small Grants program. Morgan Grant provided invaluable assistance in developing the laboratory software. We are also grateful for comments and feedback received from Peter Bossaerts, Stephen Brown, Nobuyuki Hanaki, Tibor Neugebauer, Charles Noussair, Bernt-Arne Ødegaard as well as from participants in numerous online seminars and workshops.

1 Introduction

The impact of exchange-traded funds (ETFs) on financial markets remains a subject of great debate. According to some studies (Ivanov et al., 2013; Glosten et al., 2021 and Huang et al., 2021), ETFs facilitate price discovery of the underlying assets, while other studies find that non-fundamental shocks generated in ETF markets can be transmitted to the underlying assets (Ben-David et al., 2018; Da and Shive, 2018; Brown et al., 2021 and, for the case of large ETFs, Box et al., 2021).¹ In empirical work, a key assumption when estimating the impact of ETF index inclusion on the prices of the underlying assets is that an asset that is included in the index is comparable to one that is not.² We eliminate the need for this restrictive empirical assumption by using a laboratory experiment to set up a clean counterfactual environment in which to study the impact of ETF index inclusion on asset prices. In particular, we can perfectly control whether the assets included or excluded from the index are comparable to one another.

The laboratory allows for control over other confounding variables that may also affect asset returns. For example, the literature suggests that index inclusion can create a signalling effect (Denis et al., 2003, Dhillon and Johnson, 1991 and Jain, 1987), which is then reflected in asset prices. In our study, we eliminate such signalling explanations as index membership is orthogonal to the size or the growth potential of any asset. Further, there is evidence that index inclusion alters the corporate decision-making by managers. For instance, Bennett et al. (2020) argues that when new companies are included in an index, they adopt corporate financial policies that mimic those of other included firms, which negatively affects future returns.

We study index inclusion/exclusion using an ETF index fund. The ETF market has two levels: (i) the primary market, and (ii) the secondary market. The former is open only to Authorized Participants (APs), who can create and redeem ETF shares. The latter, which accounts for about 90% of overall trading volume, is open to all investors (ICI, 2019). In our experimental design, we relegate the role of APs to robot traders or “bots” who enforce the law of one price between the ETF and the underlying assets.³

¹Box et al. (2021) find that the transmission of shocks from ETFs to the underlying assets only occurs when ETFs are large. For other classes of ETFs, there is no evidence of noise transmission.

²The natural experiment of index reconstitution of the Russell indices helps to address some of the endogeneity issues in the field work; see the methodology discussion by Appel et al. (2020) and Ben-David et al. (2019).

³Box et al. (2021) document that arbitrage opportunities rarely exist in their sample.

Our market includes three risky assets A , B and C , whose final payoff is determined by the state of nature — an Arrow-Debreu setup similar to that of [Plott and Sunder, 1982](#), [Bossaerts and Plott, 2004](#), and many others. The baseline environment, which we refer to as treatment ABC , includes an ETF asset D which is constructed using the market portfolio, or one unit of each risky asset. The second environment, the treatment labeled $A2C$, excludes asset B from the index and replaces it with a second unit of asset C . In our design, asset C is always identical to asset B in terms of fundamental value. Therefore, absent an index premium, assets B and C should have identical prices whether or not the ETF asset excludes B . Finding a discrepancy in the prices of these two assets between the two environments therefore is direct evidence of an index premium.

Specifically, we define the ETF index premium as the difference-in-difference in the price of assets C and B across environments $A2C$ and ABC . This definition allows us to generalize our findings to a broader class of subject preferences since two identical assets in a given market should have the same price; otherwise arbitrage opportunities arise. In a third environment, treatment $A2C_{short}$, we study whether relaxing the limits of arbitrage by allowing market participants to engage in short-selling of the underlying assets can reduce the ETF index premium.

We find a substantial and significant ETF index premium which persists even when short-selling is allowed. Furthermore, we find that order imbalance — the number of outstanding bids relative to asks for a given asset — is higher for the asset that is included in the index fund compared to the imbalance of the asset that is excluded from the index. Increasingly positive order imbalance reflects greater demand for assets that are included in the index. We also find that there are few arbitrage opportunities between the ETF and its underlying assets across all of our treatments. Thus, the AP-bot player’s profit for creating and redeeming the ETF share is small.

In our experiment, the AP-bot posts an ETF ask order (creating an ETF share) by reading the best ask orders of the underlying assets in the market, depending on the composition of the ETF asset, either ABC or $A2C$. A market participant can purchase the created ETF share if her bid crosses the posted ask. When the bot sells an ETF share, it almost immediately offsets the position (with latency of about 5 to 10 milliseconds) by accepting best ask orders of the underlying assets. Consequently, the bot is constrained to maintain a zero net supply. This process prevents losses by the bot and allows it to earn positive profits to the extent that ETF bid orders from

market participants are higher than the bot's ETF ask orders. In addition to creating ETF shares, the bot can also destroy ETF shares by reading the best bid orders for the underlying assets, posting a corresponding buy limit order for the ETF and, when that order is accepted, offsetting it immediately by accepting the best bid orders for the underlying assets. A market participant who is not a bot is only allowed to sell the ETF asset if she holds shares of the ETF in her portfolio. That is, across all three of our treatments, participants are not allowed to create or destroy ETF shares; this role is reserved solely for the bot who serves as the AP. The bot can earn profits when the purchase price of ETF is less than the income received from the sale of the underlying assets.

Our paper is related to the literature studying the impact of index inclusion and exclusion on stock returns. These studies show heterogeneous results. The index inclusion premium was first documented by [Shleifer \(1986\)](#), who conducted an event study of new stock inclusions to the S&P 500 Index between 1966 and 1983. He hypothesized that if the demand curve for stocks was downward sloping (and not horizontal as is typically assumed) then the decision to include a stock in an index would result in an increase in its price, as index funds would increase demand for the newly included stock causing the demand curve for that stock to shift outward. The results show that after 1976, stocks which were included in the S&P 500 index gained, on average, 2.8% the day the inclusion was announced, and that these price increases were persistent. In a later event study, [Beneish and Whaley \(1996\)](#) found that, from October 1989 through June 1994, stocks included in the S&P 500 had an average increase of 3.1% in their opening price the day following the announcement. A more recent event study, by [Petajisto \(2011\)](#), looked at the effect of inclusion and exclusion in the S&P 500 (and the Russell 2000) and found that there was an average 8.8% (4.7%) increase in the price of a newly included stock from the date of the announcement to the actual inclusion date. The effect of exclusion from the S&P 500 (Russell 2000) led to a decrease of 15.1% (4.6%) following the exclusion announcement. However, index premiums have been decreasing since the 2000s. Using a regression discontinuity approach on the Russell 2000, [Chang et al. \(2015\)](#) find that addition to the index results in a price increase, while deletion leads to a price declines. Our paper is the first to provide evidence for an ETF index inclusion premium using a clean counterfactual environment.

2 Environment

In our Arrow-Debreu environment, there are N market participants who trade three risky assets, $\theta \in \{A, B, C\}$, and an index ETF asset, whose composition is known to all participants: (1) $ETF = A + B + C$ in the ABC treatment, or (2) $ETF = A + 2C$ in the $A2C$ treatment. All market participants are endowed with some combination of risky assets θ , as described in Table 1, and a risk-free loan L . The market supply of each asset, θ , is two units per capita, and the ETF is constrained to be in zero net supply. A third treatment, $A2C_{short}$, is the same as treatment $A2C$ except that subjects in treatment $A2C_{short}$ can short sell each of the θ assets.

Table 1: Endowment per player 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 two types of market participants: type I and type II. Regardless of type, each market participant receives a loan equal to 210.

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 asset A (80) and assets 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 assets B and C yield the same terminal payoff. Since the ETF asset is offered at zero net supply, we can use the standard asset pricing model to predict asset prices. However, in our case, point price predictions are not of interest; rather, we focus on relative prices of identical assets to evaluate the impact of index inclusion.⁴

At the end of each trading period, traders' 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,

⁴To provide exact asset pricing predictions, we need to make assumptions about utility functions and the degree of risk aversion. For the interested reader, more detailed information on asset pricing predictions can be found in [Bossaerts and Plott \(2004\)](#) or [Bossaerts et al. \(2007\)](#).

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 ETF composition is either $A + B + C$ or $A + 2 \times C$.

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 also be covered by subjects within each trading period.

As a market clearing mechanism, we employ a continuous double auction (CDA) for each asset. Specifically, market participants can trade in up to four separate markets simultaneously, with one market assigned to each asset. In the CDA format, traders can participate in any market by either (i) submitting a limit order to buy (a bid) or to sell (an ask) for a single unit, or (ii) 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 on price, 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 trader accepts an existing order by clicking on it and agreeing to the buy or sale price. The transaction price is always that of the earlier-in-time order. As previously noted, in treatment $A2C_{short}$, market participants are allowed to short-sell each of the three risky assets θ , but not the ETF.

2.1 Authorized participants

In 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 ETF shares outstanding to keep the price of the ETF aligned with the value of its underlying assets. In our setting, the AP 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) 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 market participants 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 underlying 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 , such that $a_{ETF} = a_A + a_B + a_C$. If a market participant purchases the ETF offered by the AP-bot, then the bot will immediately accept the outstanding asks of the assets A, B , and C . Similarly, if the AP-bot finds outstanding bid orders for all of the underlying assets, then it will post a bid order for the ETF , e.g. $b_{ETF} = b_A + b_B + b_C$. If a market participant accepts that bid, then the AP-bot will immediately accept the bids for the underlying assets.

2.2 Hypotheses

The prices for assets B and C should be equivalent within a treatment and across treatments since they have the same payoff structure; the endowment of assets does not vary across treatments, and the ETF is offered in zero net supply. However, if ETF index inclusion generates an additional value or premium, then we will observe a divergence in the prices of the two equivalent assets. Below, we formulate our hypotheses based on the idea that ETF index inclusion creates a price premium for the included asset.

Hypothesis 1: *The price of asset C will be greater than the price of asset B in treatments $A2C$, and $A2C_{short}$, indicating a positive index premium φ_I .*

To test for the ETF index premium, we compare the price difference between equivalent assets B and C in treatment ABC against the price difference in treatment $A2C$. The price difference between the two assets in a treatment M can be written as $\varphi_P^M := P_C - P_B$. Thus, we define the ETF index (I) premium as

$$\varphi_I := \varphi_P^{A2C} - \varphi_P^{ABC}, \quad (1)$$

and for the treatment $A2C_{short}$

$$\varphi_I^{short} := \varphi_P^{A2C_{short}} - \varphi_P^{ABC}. \quad (2)$$

Given that there are no differences in the characteristics of the two assets, B and C , any price difference between those two assets can be only attributed to ETF index inclusion. In treatment ABC we expect that $\varphi_P^{ABC} = 0$, and in treatment $A2C$, $\varphi_P^{A2C} > 0$. Therefore, the ETF index premium φ_I will be positive in equation (1). Although short selling may have some effect on prices, we hypothesize that it will not make the index premium disappear. Hence we predict that $\varphi_I^{short} > 0$ in equation (2).

Hypothesis 2: *If there is a positive ETF index inclusion premium, then order imbalance will be positive for assets included in the ETF index and negative for assets excluded from the ETF index.*

To explain, Hypothesis 2 let $d \in (0, 1]$ be a weight parameter which discounts orders more heavily the further they are away from the midpoint $m(t)$ between the current best bid and ask prices, and let $Q(p, t) \in Z$ be the number of outstanding + sell (– buy) orders at price p at time t . Following the literature, we then define the order imbalance as

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

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,

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

and

$$\varphi_O^{short} := \varphi_z^{A2C_{short}} - \varphi_z^{ABC}. \quad (5)$$

Preference for index products, and its constituents, should lead to an increase in the demand for asset C , and a decrease in the demand for asset B in treatment $A2C$. The increase in demand should be related the number of bids relative to asks in the order book. For asset C , we expect to observe more bids relative to asks, and for asset B we expect to observe less bids relative to asks. For asset B , which is excluded from the index in treatments $A2C$ and $A2C_{short}$, the greater number of asks relative to bids should lead to an increasingly negative order imbalance.

Thus, we expect that $z_B < z_C$, which results in $\varphi_z^{A2C} > 0$ and $\varphi_z^{A2C_{short}} > 0$. In the treatment ABC , where assets B and C are both included in the index, we do not expect a significant value for φ_z^{ABC} . Therefore, we expect that $\varphi_O, \varphi_O^{short} > 0$.

Hypothesis 3: *Short-selling decreases the index premium, and order imbalance.*

When short-selling is permitted, as in our $A2C_{short}$ treatment, we expect the price differential between asset B and asset C to decrease because arbitrage is facilitated. Therefore, allowing market participants to engage in short-selling should decrease the ETF index premium, such that

$$\varphi_I^{short} < \varphi_I, \quad (6)$$

and the order imbalance,

$$\varphi_O^{short} < \varphi_O. \quad (7)$$

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. Participants 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 Online Appendix. 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. In total, we conducted 15 sessions, with 5 sessions

for each of our 3 treatments, using between 10 and 16 subjects per session. Average earnings are \$13.95 which excludes a show-up fee of \$7.⁵ We present an overview of all sessions from our experiment in Table 3.

Table 3: Overview of experimental sessions

Treatment	Sessions	Participants	Earnings (\$, mean)
<i>ABC</i>	5	58	12.81
<i>A2C</i>	5	62	16.40
<i>A2C_{short}</i>	5	62	12.57
Total	15	182	13.95

Note: Each session had 10-16 participants. The participants also received \$7 as show-up fee.

3.1 Trading interface

Figure 1 shows the market user interface (UI), which was developed in oTree the software used to conduct the experiment. The top panel of the UI is divided into four quadrants, where each quadrant represents an asset market. For each asset market, a participant can view the order book with outstanding Bids and Asks, and all past traded prices (Trades). The bottom panel of the UI is also divided into four parts and provides the following information: (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.

Market participants can submit single limit orders by typing a buy (sell) price in the corresponding input box for each asset and clicking on the button labeled “Buy” (“Sell”). The order submitted by a participant is immediately posted to the order book for the relevant asset and is identified by a red cross next to the order. The participant can cancel their order at any time while it remains active 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 can 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 earlier-in-time limit order bid (ask). Second, market buy or sell orders can be

⁵Higher earnings in treatment *A2C* are due to state *Z* being realized in periods randomly drawn for payment. Recall that state *Z* yields a higher asset payoff.

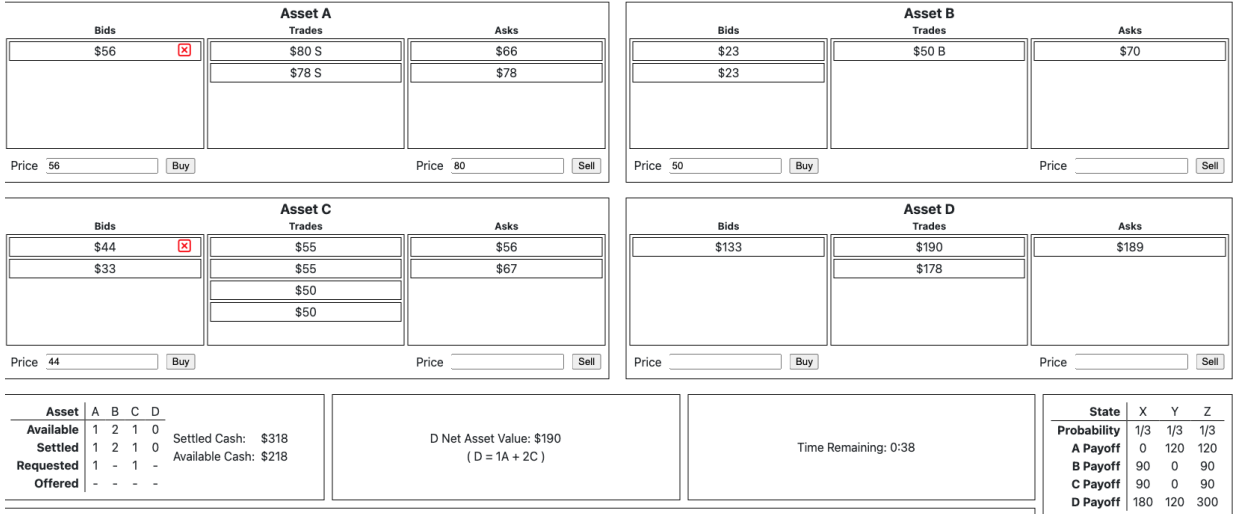


Figure 1: User interface (UI) $A2C_{short}$ treatment; other interfaces are similar.

Note: The red cross denotes the participant order, and by clicking on it she can cancel the order. When a participant traded, the letter S next to the order means that she sold, and B, bought. The AP-bot in this case is posting asks and bids for the ETF, asset D which is equal to the sum of the bids (asks) of the underlying assets, which in this case is a unit of A and two units of C.

directly executed by clicking on any existing bids (asks) in any of the simultaneous four asset market order books and confirming the intention to buy (sell) at that price which then generates a trade between participants. All transactions are recorded and presented in the center of the order book under the heading “Trades”. A letter “B” or “S” indicates that the transaction belongs to the participant viewing the book, and whether she bought (B) or sold (S) the single asset. For all treatments, the bids submitted by a participant cannot exceed their 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 If the final position in any asset remains negative, then the participant must buy back the asset at its terminal value, as determined by the state, s . Asset D , the *ETF* asset, is a composite asset that is formed as a combination of the θ assets depending on the treatment. At the start of each trading period, none of the participants hold any D assets, but the AP-bot may offer units of asset D for sale following the start of trade in each period according to the rules discussed earlier; there has to be ask orders for all assets included in the ETF for the bot to offer an ETF asset D to the market. Human traders can submit bids for D at any point in

⁶A human trader 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 should be sufficient to reduce arbitrage opportunities. Of course, there is also a limit on the cash available in the market which constrains the asset turnover.

the market, but they can submit asks for D if and 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 sessions: (i) a representative example session from treatment ABC as shown in Figure 2, and (ii) a representative example session from treatment $A2C$ as shown in Figure 3. A complete depiction of results from each and every session can be found in the Appendix.

In the ABC treatment session shown in Figure 2, the top panel presents prices for all assets. We see that the prices of assets B and C oscillate around the expected value (the dashed gray line at 60). The prices for the ETF and asset A are higher, and approach the expected value as depicted by 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 imbalance for the two identical assets B (red), and C (orange) are generally the same.⁷ 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 asset B is excluded from the index and is replaced by asset C . The top panel, which presents prices, shows that the price of asset B (red) is below the price of asset C (orange). Despite being identical, market participants appear to place a high value on asset C that is included in the index by trading it at a higher price relative to the asset that is excluded, asset B . Recall that we do not observe this divergence in Figure 2. Given that the only difference between treatments ABC and $A2C$ is the composition of the ETF asset, namely the replacement of asset B by asset C in the index, the price differential should be due to the ETF index premium. Thus, being included in the index appears to lead to an increase in the asset price.

The divergence in the prices of assets B and C in the $A2C$ treatment is further analyzed by comparing the order imbalances between the two assets. The bottom panel of Figure 3 shows that the order imbalance for asset B (red line) is consistently

⁷Our measure of order imbalance in equation (3) assumes a value of $d = 0.99$ with the objective to only discount orders that are extreme outliers.

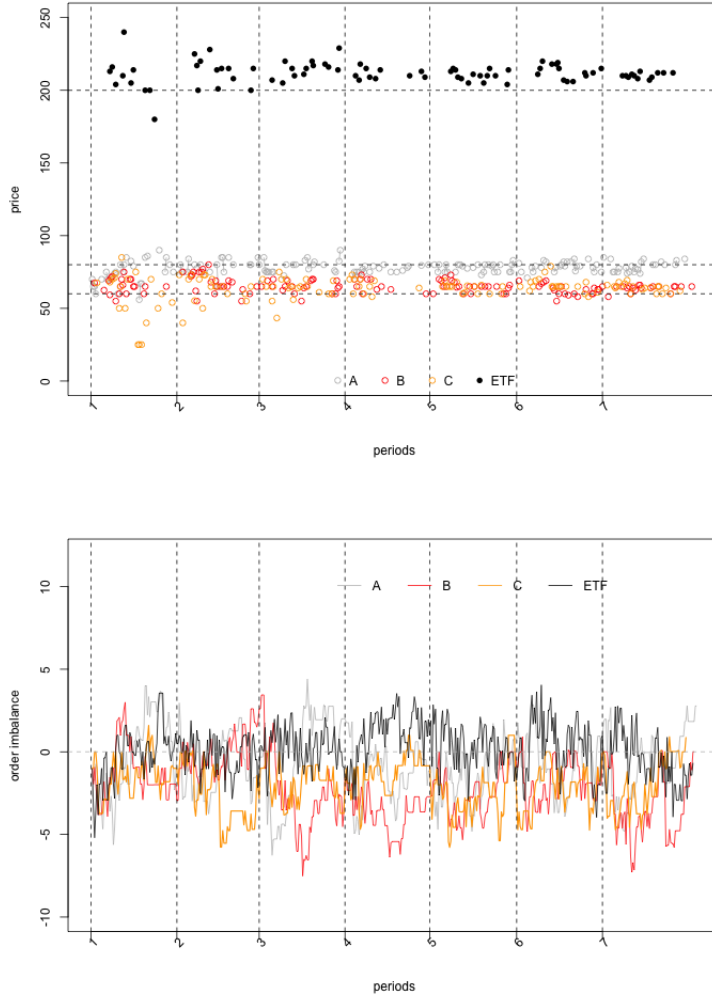


Figure 2: Prices and order imbalance for *ABC* treatment (a session example)

Note: The top (bottom) panel presents asset 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 each asset: 200, 80 and 60 for the ETF, *A*, and *B* or *C*, respectively.

below the order imbalance for asset *C* (orange line). This suggests that when an identical asset is excluded from the index (*B*), it will have an increasingly negative order imbalance (i.e., there are more outstanding asks than bids), relative to an asset that is included in the index (*C*). Thus, the idea of a price premium due to index inclusion is supported by a close examination of the order book, which shows that the included asset will feature more bid offers, relative to the excluded asset.

Figure 4 offers a useful data summary, of period-by-period average price differences (left panel) and the average order imbalances differences (right panel) between assets *B*

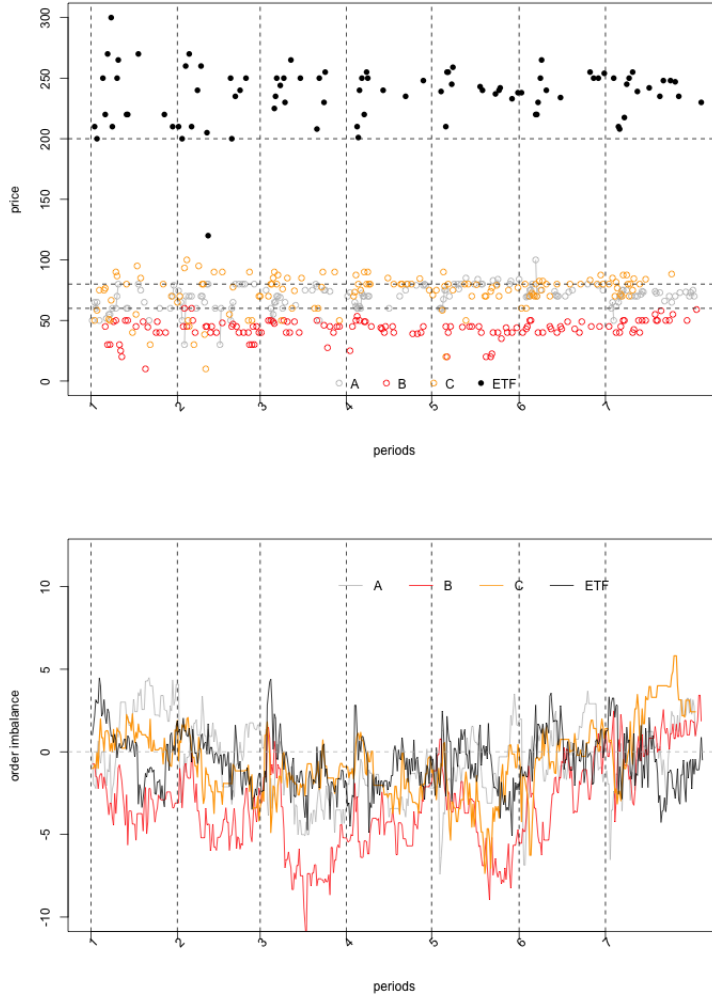


Figure 3: Prices and order imbalance for $A2C$ treatment (a session example)

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.

and C . Black lines denote the ABC treatment, blue dashed lines the $A2C$ treatment, and red dashed lines the $A2C_{short}$ treatment. Consistent with Hypothesis 1, 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, however, the prices of the two identical assets appear to be the same; the price differential is close to zero. Also, consistent with Hypothesis 2, the mean difference between order imbalances across assets B and C (of about 2) for treatments $A2C$ and $A2C_{short}$ is generally greater than the one observed in treatment ABC , whose value is usually close to zero. The

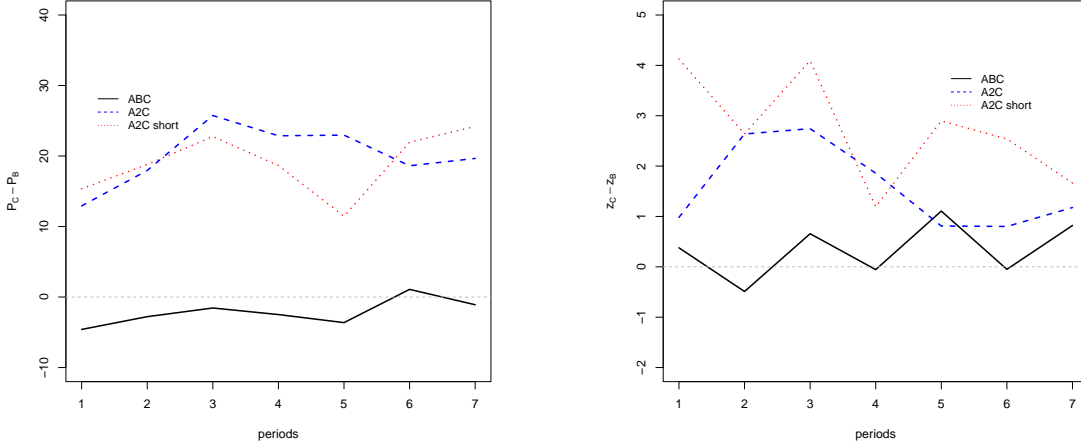


Figure 4: Mean price difference (left) and order imbalance (right) for assets B and C
Note: Each session includes 7 periods, and means over the 5 sessions are shown for each of the 3 treatments, ABC , $A2C$ and $A2C_{short}$.

greatest order imbalance is exhibited in $A2C_{short}$, the treatment that permits short-selling.

4.1 Hypothesis Test Results

Result 1: *The ETF index premium φ_I is positive in treatments $A2C$, and $A2C_{short}$. When asset B is excluded from the index it is priced lower than asset C .*

According to Figure 4, the price difference $\varphi_P^M := P_C - P_B$ is close to zero across all seven periods in treatment ABC , but is greater than zero in treatments $A2C$ and $A2C_{short}$. To confirm the existence of an ETF index premium, we report an OLS regression in Table 4. The regression includes fixed effects at the market level and standard errors that are clustered at the session level. In the first column, the dependent variable is the asset price, and the independent variables include: dummies for assets A , C and for the *ETF*; treatment dummies $A2C$ and $A2C_{short}$; and finally interaction terms between assets and treatments. The constant term thus represents the price of asset B in treatment ABC .

Our regression analysis shows that the prices of assets B and C in the ABC treatment are very similar. There is a statistical significance (p -value < 0.001) for the coefficient associated to asset C , but the difference is small (about 2, and asset B is the one that is priced higher). In treatment $A2C$, the price of asset B decreases by about 13 (p -value of 0.038) relative to treatment ABC . Asset C is priced about

Table 4: Prices and order imbalance (OLS)

	Price	p -value	Order Imbalance	p -value
constant	64.65	0.000	-2.06	0.000
A	12.15	0.000	1.85	0.000
C	-2.24	0.000	0.33	0.035
ETF	148.46	0.000	2.45	0.000
A2C	-13.42	0.038	-0.20	0.844
A \times A2C	11.15	0.008	-0.28	0.651
C \times A2C	22.17	0.000	1.25	0.011
ETF \times A2C	41.78	0.000	0.15	0.836
A2C _{short}	-24.70	0.001	-2.29	0.052
A \times A2C _{short}	8.04	0.081	2.46	0.007
C \times A2C _{short}	21.21	0.000	2.40	0.000
ETF \times A2C _{short}	9.94	0.648	2.66	0.001
R^2	0.88		0.24	
N	4 960		30 264	

Notes: Order imbalance is defined by equation (3). The regression includes fixed effects at the market level, and the standard errors are clustered at the session level. The variable $A2C$ takes the value of one if the ETF is constructed as $A + 2C$, and zero otherwise. The variable $A2C_{short}$ takes the value of one if the ETF is constructed as in the environment $A2C$ and short-selling is allowed, and zero otherwise.

8.75 = 22.17 - 13.42 points higher than in treatment ABC , according to the interaction term $C \times A2C$ and the treatment dummy $A2C$. However, this is not statistically different than zero (p -value of 0.09 using a Wald test). The ETF index premium, which is captured by the coefficient $C \times A2C$, is about 22 and strongly significant (p -value < 0.001).

The price of asset B in treatment $A2C_{short}$ is also lower (by about 25) compared to the treatment ABC (p -value of 0.001). Furthermore, the price of asset C is not statistically different from its price in treatment ABC (p -value of 0.70 using a Wald test), captured by the interaction term $C \times A2C_{short}$ and the treatment dummy $A2C_{short}$. The ETF index premium, measured by the coefficient of $C \times A2C_{short}$, is 21.21 (p -value < 0.001). Thus, we confirm our prediction in Hypothesis 1, and find evidence for a significantly positive ETF index premium. We conclude that in treatments where an identical asset is excluded from the index, identical assets will be priced differently. In our data, the value of ETF index inclusion is about a third of the expected payoff for the excluded asset.

Result 2: *When an identical asset is excluded from the index, the order imbalance becomes increasingly positive for the included asset.*

To analyze the impact of exclusion, the last columns in Table 4 report a parallel regression whose dependent variable is the order imbalance defined in equation (3). The constant in the regression captures the order imbalance for asset B in the ABC treatment; it is -2.06 (p -value < 0.001). For asset A and the ETF we cannot reject that order imbalance is equal to zero (p -value of 0.64 and 0.41, respectively, using a Wald test), however for asset C , the order imbalance is also negative ($0.33 - 2.06$, with a p -value of 0.006 using a Wald test). A negative order imbalance implies that the order book for assets B and C has more ask orders than bid orders.

More importantly, the impact of ETF index inclusion in the $A2C$ treatment is not significant, which implies that the order imbalance for asset B is not different across treatments $A2C$ and ABC . However, the interaction term $C \times A2C$ of 1.25 shows a positive order imbalance premium (p -value of 0.011), which means that in the $A2C$ treatment the order book for asset C has more bid orders than asks orders compared to asset B , indicating increasing demand for asset C .

The impact of ETF index inclusion on the order imbalance appears even stronger in the $A2C_{short}$ treatment. The interaction term $C \times A2C_{short}$ is 2.40 (p -value < 0.001), which can be explained by the larger number of underlying assets available for trade when the short-selling constraint is relaxed.

Result 3: *The ETF index premium and order imbalance is the same across the $A2C$ and the $A2C_{short}$ treatments.*

We formally test whether the index premium and the order imbalances differ across $A2C$ and the $A2C_{short}$ by performing a Wald test to determine whether the coefficients for $C \times A2C$ and $C \times A2C_{short}$ are equal for both regression specifications. The test suggests that we cannot reject that lifting short-selling constraints leads to the same outcome as in the constrained environment (p -value of 0.877 for prices, and 0.055 for order imbalance).

4.2 Other findings

Result 4: *Asset turnover for the asset excluded from the index is significantly lower than for the identical asset that is included in the index in treatment $A2C$.*

Next, we study how asset turnover is affected when an identical asset is excluded from the index. Table 5 presents the asset turnover for assets A , B , and the ETF relative to asset C .⁸ In treatment ABC , the asset turnover for the two identical assets, B and C , is one to one. However, when asset B is excluded from the index, in treatments $A2C$ and $A2C_{short}$, its asset turnover is lower, at 0.65 and 0.71 units respectively. We formally test the difference of asset turnovers across treatments, and find statistically significant differences only for the treatment $A2C$ (p -value of 0.008 using a Wilcoxon test). We conclude that short-selling constraints decrease trading activity in the excluded asset, B .

Table 5: Asset turnover (in terms of asset C)

Treatment	A	B	ETF
ABC	1.06	1.00	0.43
$A2C$	0.71	0.65	0.31
$A2C_{short}$	0.92	0.71	0.30

Note: For each session, we divide the total number of units transacted, for assets A , B and the ETF , by the total number of units traded of asset C , and present the average turnover per treatment. The supply per capita, across all sessions, is two units of assets A , B and C , respectively.

Result 5: *The AP-bot trader generates only a small profit from trading activity.*

Since the role of the AP-bot trader is to enforce the law of one price, most of the transactions occur at the price where the ETF is equal to the market value of the underlying assets or the NAV . As explained earlier, the AP-bot can earn a profit when it sells (buys) the ETF at a higher (lower) price than the NAV , whose composition depends on the treatment conditions. In Table 6 we analyze the average AP-bot profit and arbitrage activity across our three treatments.

The first column of Table 6 presents the average profit: it appears that bot's 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 . The AP-bot's profit per transaction is quite

⁸Asset turnover is defined as the total number of units transacted. We work in terms of units of asset C to control for the different number of subjects across sessions, and emphasize the difference with respect to the excluded asset B in some treatments.

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 $A2C$.

Table 6: 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.

5 Conclusion

Consistent with the *downward sloping demand* explanation of Shleifer (1986), we find a substantial and significant index premium in a lab setting that excludes other channels, such as signalling, changes in corporate governance and/or investor awareness. Moreover, we find that the order imbalance that arises when an identical asset is excluded from the index, plays a significant role in explaining the observed index premium. Relaxing the short-selling constraint for the underlying assets does not eliminate the index premium.

In the field, Pavlova and Sikorskaya (2022) also find a downward sloping demand curve that is derived from both passive and active investors, where the latter use index funds as a benchmark. In the experimental literature, there is little work on the consequences of index products in asset markets. The most related paper, by Duffy et al. (2021), finds that an ETF index can provide a useful benchmark, and reduce mispricing of the underlying stocks. Our results in the ABC treatment confirm these findings—when the index covers the entire market, stocks are priced according to the fundamentals. However, Duffy et al. (2021) do not consider environments in which the index asset fails to cover *all* available underlying assets. Additionally, their design does not include AP players who create and redeem ETF shares in line with the prices of the underlying assets.⁹

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

⁹For an overview of experimental work in a multiple asset environment, the interested reader can refer to Duffy et al. (2022).

Neugebauer (2019) find that the LOP holds for two assets with perfectly positive correlation. Our results show that ETF index assets may contribute to *violations* of the LOP, when they do not completely cover all the assets in a market. Future work along in this area might analyze the role of different investor types (active versus passive) or different ETF products (e.g., leverage ETFs), among other topics in which an experimental approach can shed more light on the forces affecting asset prices.

References

- Appel, Ian, Todd A Gormley, and Donald B Keim, “Identification using Russell 1000/2000 index assignments: A discussion of methodologies,” *Critical Finance Review, Forthcoming*, 2020.
- Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, “Do ETFs increase volatility?,” *The Journal of Finance*, 2018, *73* (6), 2471–2535.
- , – , and – , “An improved method to predict assignment of stocks into russell indexes,” Technical Report, National Bureau of Economic Research 2019.
- Beneish, Messod D and Robert E Whaley, “An anatomy of the “S&P Game”: The effects of changing the rules,” *The Journal of Finance*, 1996, *51* (5), 1909–1930.
- Bennett, Benjamin, René M Stulz, and Zexi Wang, “Does joining the S&P 500 index hurt firms?,” Technical Report, National Bureau of Economic Research 2020.
- 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.
- , – , and William R Zame, “Prices and portfolio choices in financial markets: Theory, econometrics, experiments,” *Econometrica*, 2007, *75* (4), 993–1038.
- Box, Travis, Ryan Davis, Richard Evans, and Andrew Lynch, “Intraday arbitrage between ETFs and their underlying portfolios,” *Journal of Financial Economics*, 2021.
- Brown, David C, Shaun William Davies, and Matthew C Ringgenberg, “ETF arbitrage, non-fundamental demand, and return predictability,” *Review of Finance*, 2021, *25* (4), 937–972.

- 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.
- Da, Zhi and Sophie Shive**, “Exchange traded funds and asset return correlations,” *European Financial Management*, 2018, 24 (1), 136–168.
- 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., *Forthcoming in Handbook of Experimental Finance*, Edward Elgar Publishing, 2022.
- Fisher, Eric O’N and Frank S Kelly**, “Experimental foreign exchange markets,” *Pacific Economic Review*, 2000, 5 (3), 365–387.
- Glosten, Lawrence, Suresh Nallareddy, and Yuan Zou**, “ETF activity and informational efficiency of underlying securities,” *Management Science*, 2021, 67 (1), 22–47.

- Huang, Shiyang, Maureen O’Hara, and Zhuo Zhong**, “Innovation and informed trading: Evidence from industry ETFs,” *The Review of Financial Studies*, 2021, *34* (3), 1280–1316.
- ICI**, *Fact Book*, Investment Company Institute, 2019.
- Ivanov, Stoyu I, Frank J Jones, and Janis K Zaima**, “Analysis of DJIA, S&P 500, S&P 400, NASDAQ 100 and Russell 2000 ETFs and their influence on price discovery,” *Global Finance Journal*, 2013, *24* (3), 171–187.
- 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.
- Pavlova, Anna and Taisiya Sikorskaya**, “Benchmarking intensity,” 2022.
- Petajisto, Antti**, “The index premium and its hidden cost for index funds,” *Journal of Empirical Finance*, 2011, *18* (2), 271–288.
- Plott, Charles R and Shyam Sunder**, “Efficiency of experimental security markets with insider information: An application of rational-expectations models,” *Journal of political economy*, 1982, *90* (4), 663–698.
- Shleifer, Andrei**, “Do demand curves for stocks slope down?,” *The Journal of Finance*, 1986, *41* (3), 579–590.

Appendix. Summary of sessions.

Sessions for treatment *ABC*

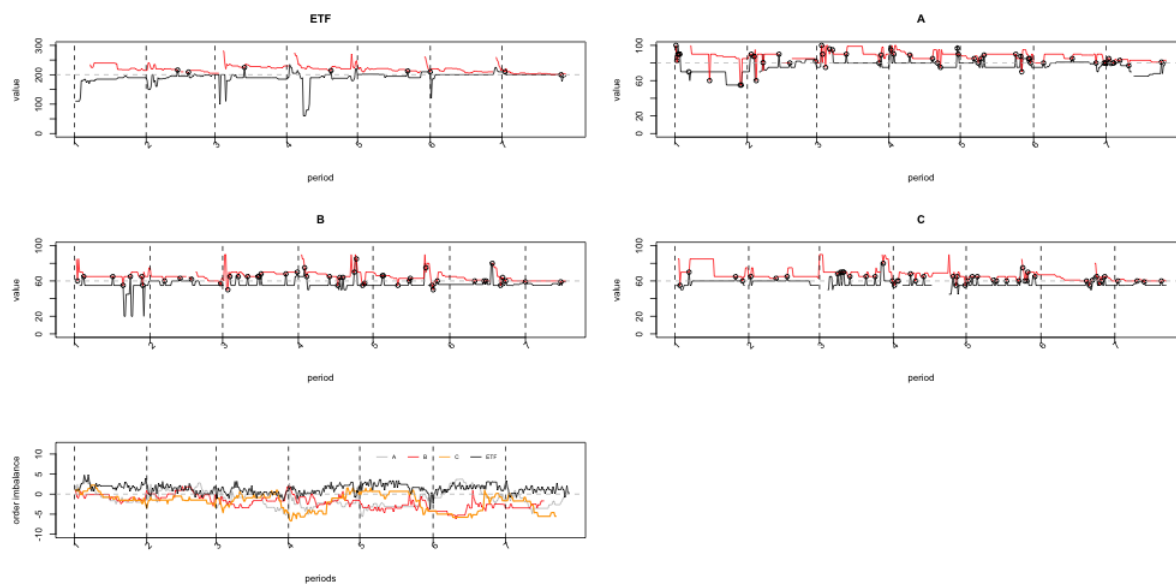


Figure 5: Best bids/offers, prices and order imbalance for *ABC* treatment (session 1)

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 asset *A* (gray), asset *B* (red), asset *C* (orange), and the *ETF* (black).

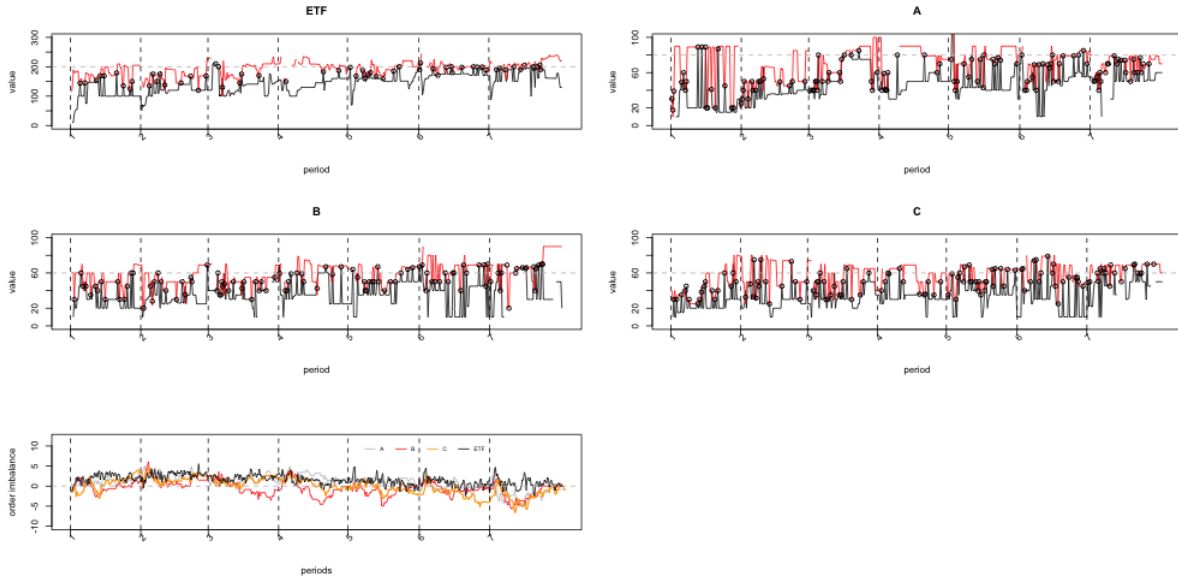


Figure 6: Best bids/offers, prices and order imbalance for *ABC* treatment (session 2)

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).

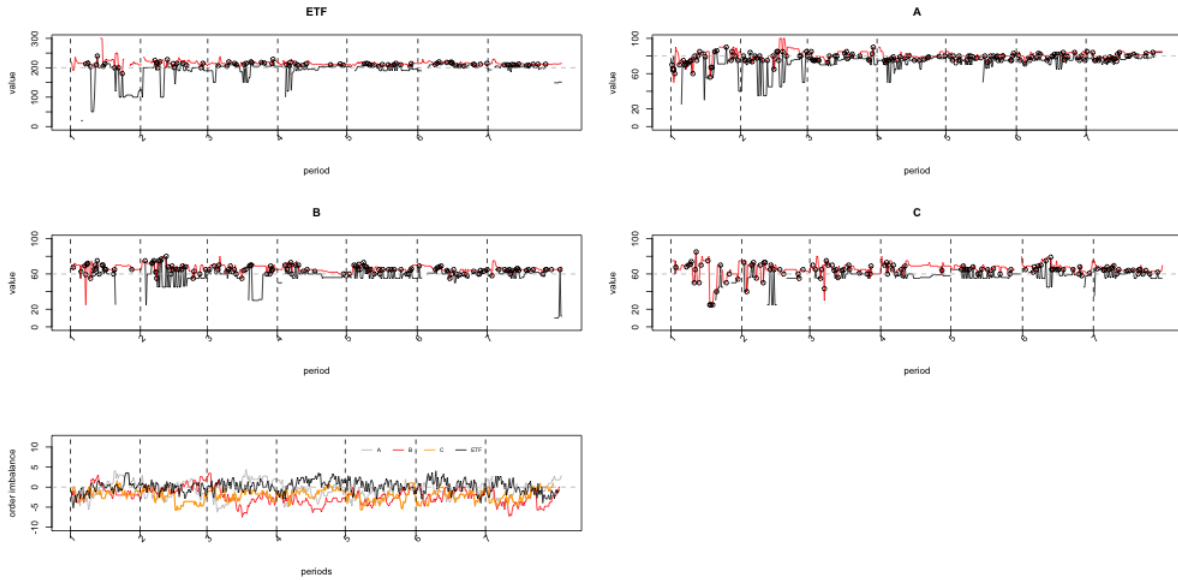


Figure 7: Best bids/offers, prices and order imbalance for *ABC* treatment (session 3)

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).

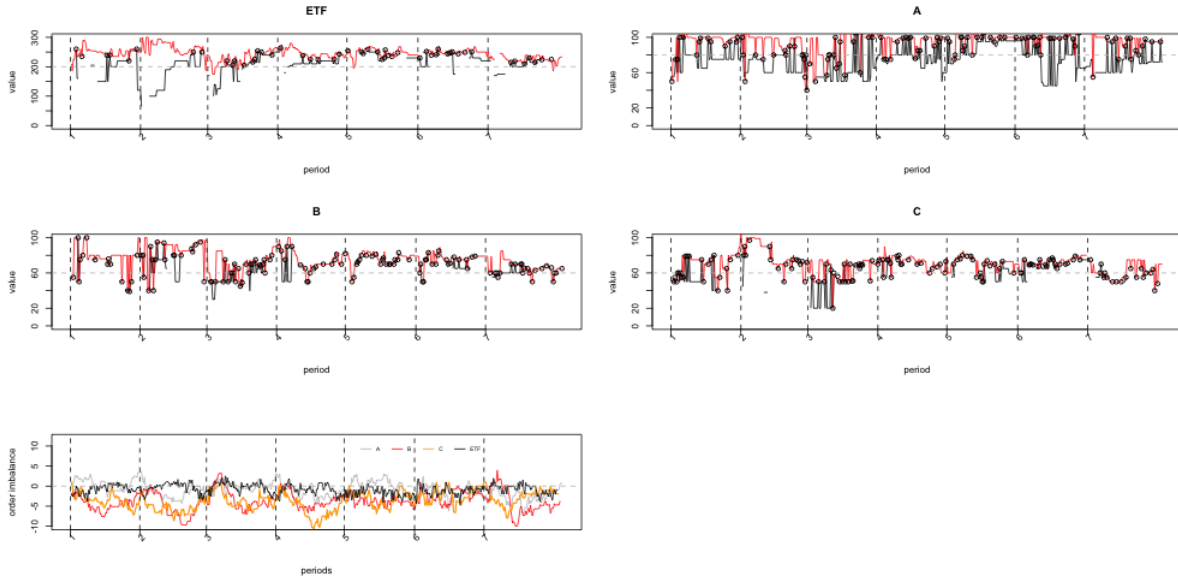


Figure 8: Best bids/offers, prices and order imbalance for *ABC* treatment (session 4)

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).

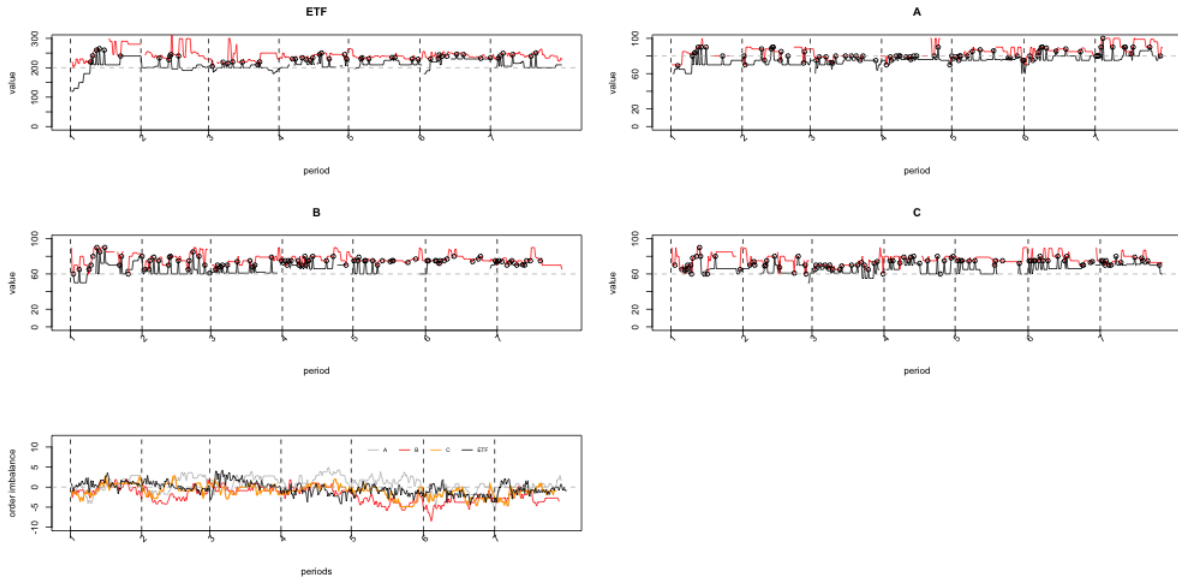


Figure 9: Best bids/offers, prices and order imbalance for *ABC* treatment (session 5)

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 $A2C$

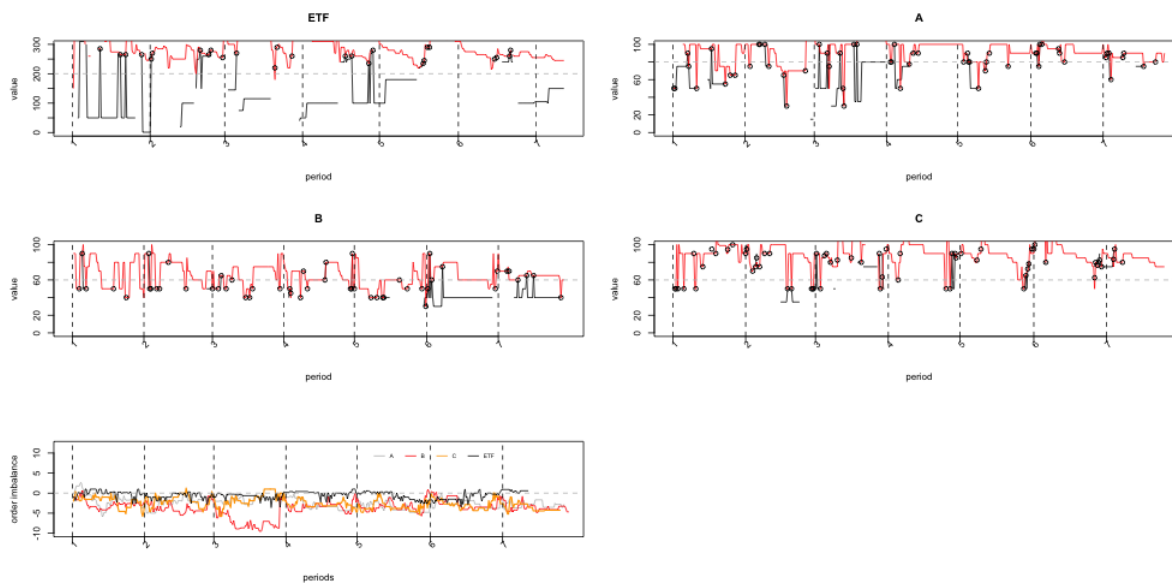


Figure 10: Best bids/offers, prices and order imbalance for $A2C$ treatment (session 1)

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).

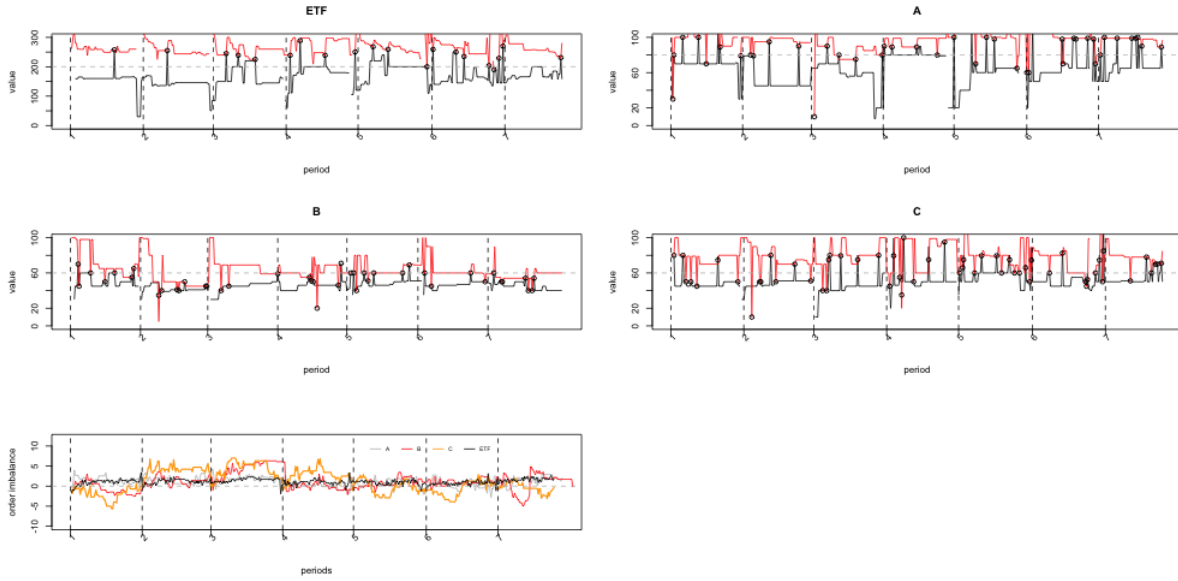


Figure 11: Best bids/offers, prices and order imbalance for $A2C$ treatment (session 2)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF , A , B and C . Each tick corresponds to four seconds, and the dashed vertical lines denote 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).

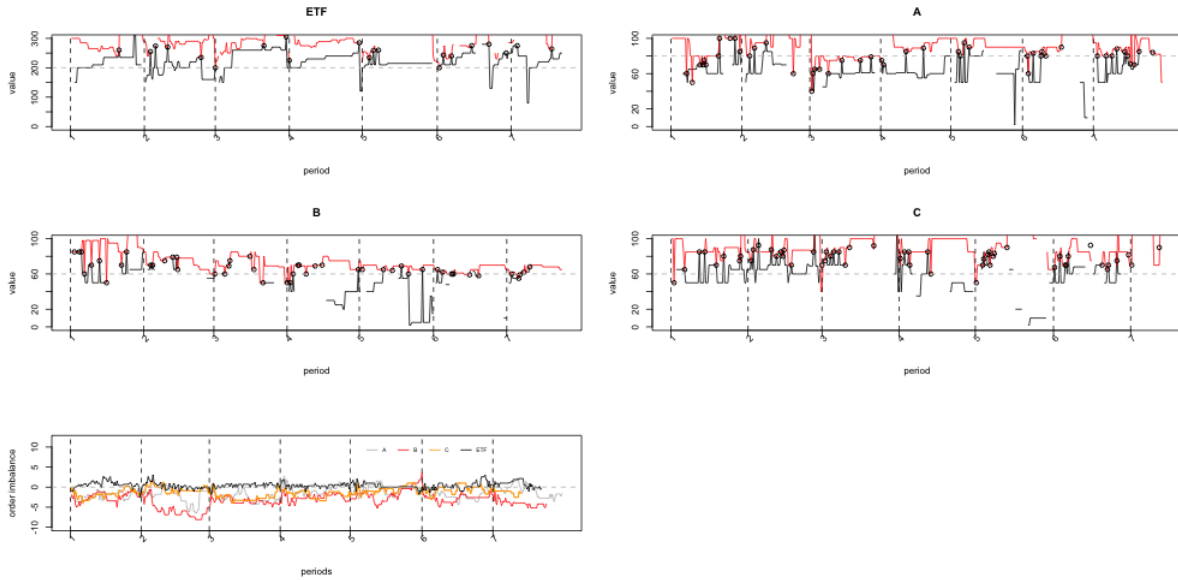


Figure 12: Best bids/offers, prices and order imbalance for $A2C$ treatment (session 3)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF , A , B and C . Each tick corresponds to four seconds, and the dashed vertical lines denote 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).

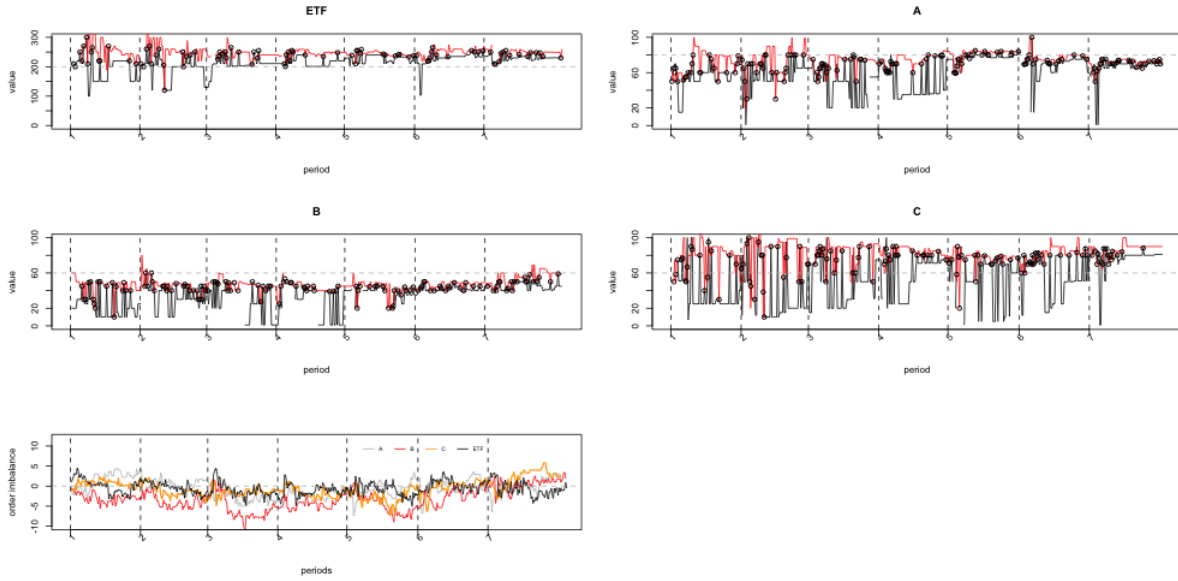


Figure 13: Best bids/offers, prices and order imbalance for $A2C$ treatment (session 4)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF , A , B and C . Each tick corresponds to four seconds, and the dashed vertical lines denote 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).

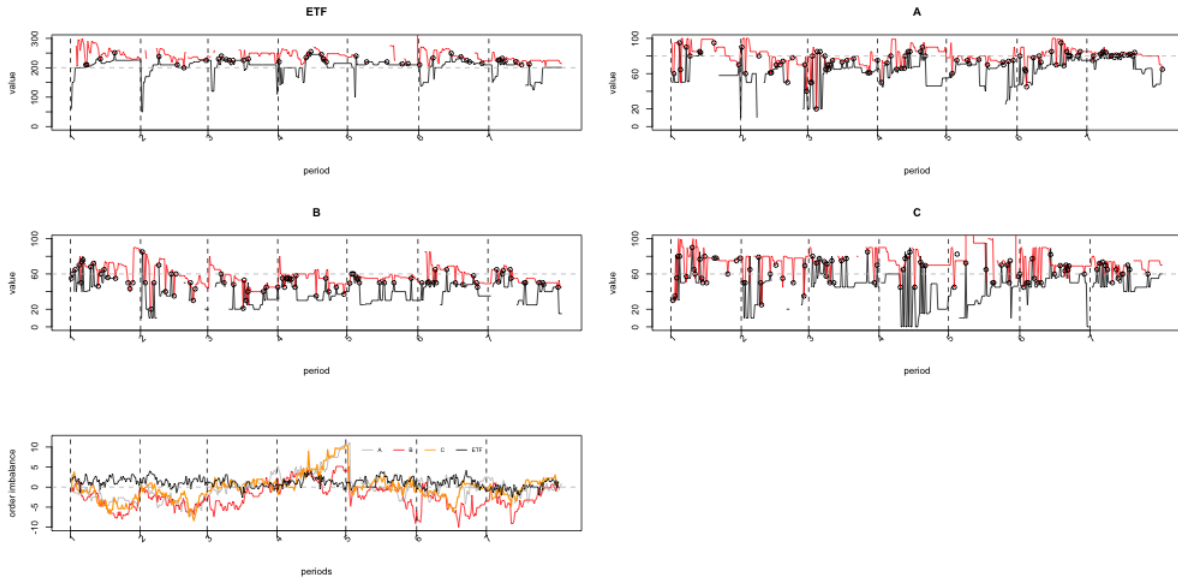


Figure 14: Best bids/offers, prices and order imbalance for $A2C$ treatment (session 5)
Note: The first two rows present the best bid (black), best offer (red) and price (dots) for each of the four markets: ETF , A , B and C . Each tick corresponds to four seconds, and the dashed vertical lines denote 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 $A2C_{short}$

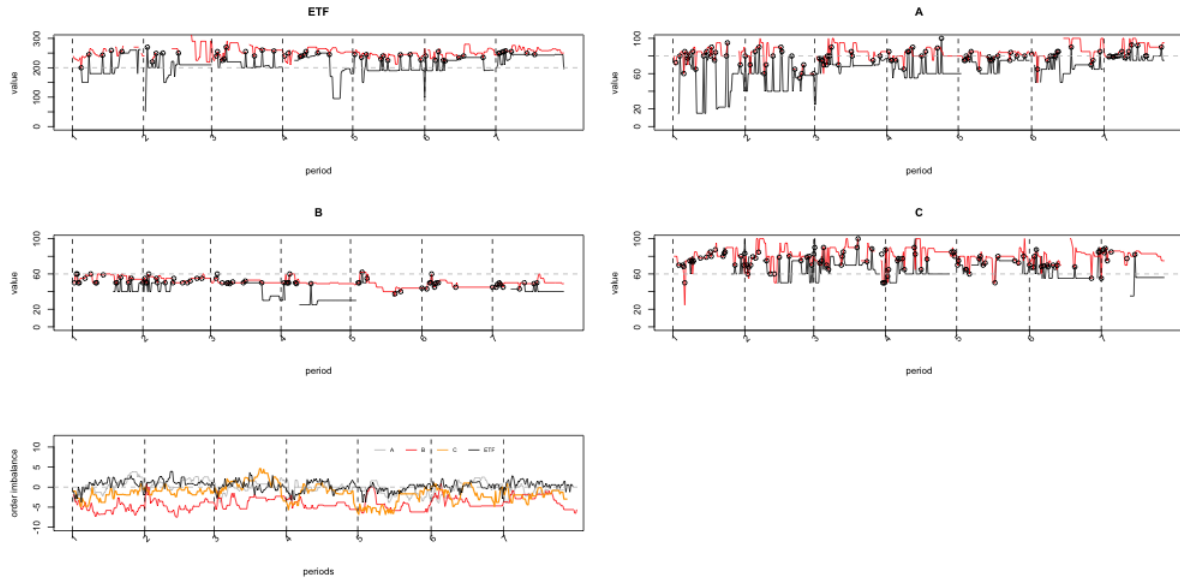


Figure 15: Best bids/offers, prices and order imbalance for $A2C_{short}$ treatment (session 1)

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).

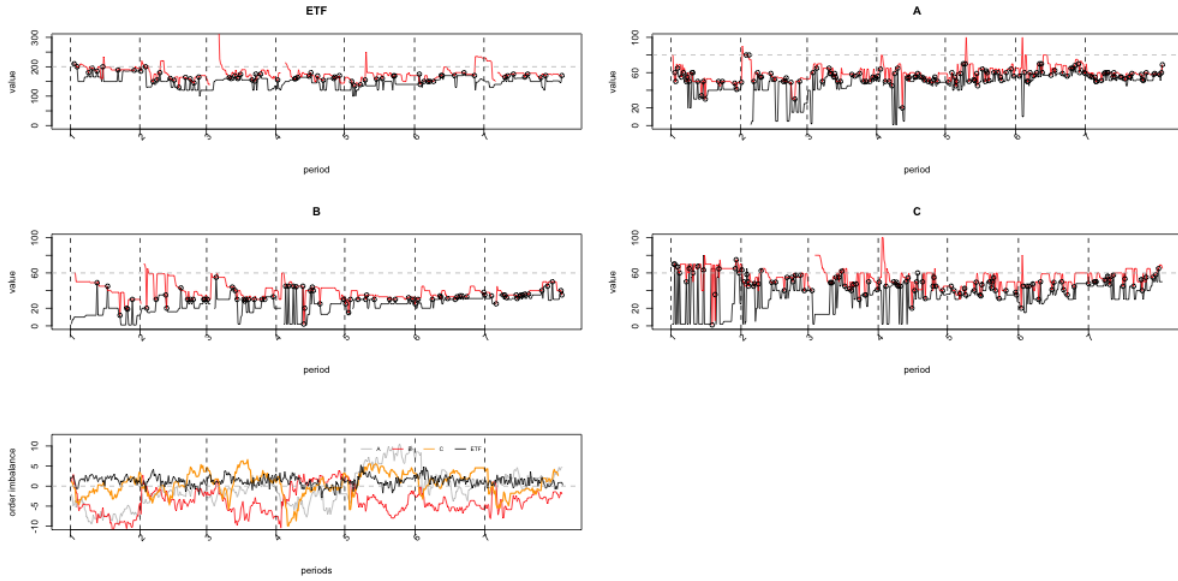


Figure 16: Best bids/offers, prices and order imbalance for $A2C_{short}$ treatment (session 2)

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).

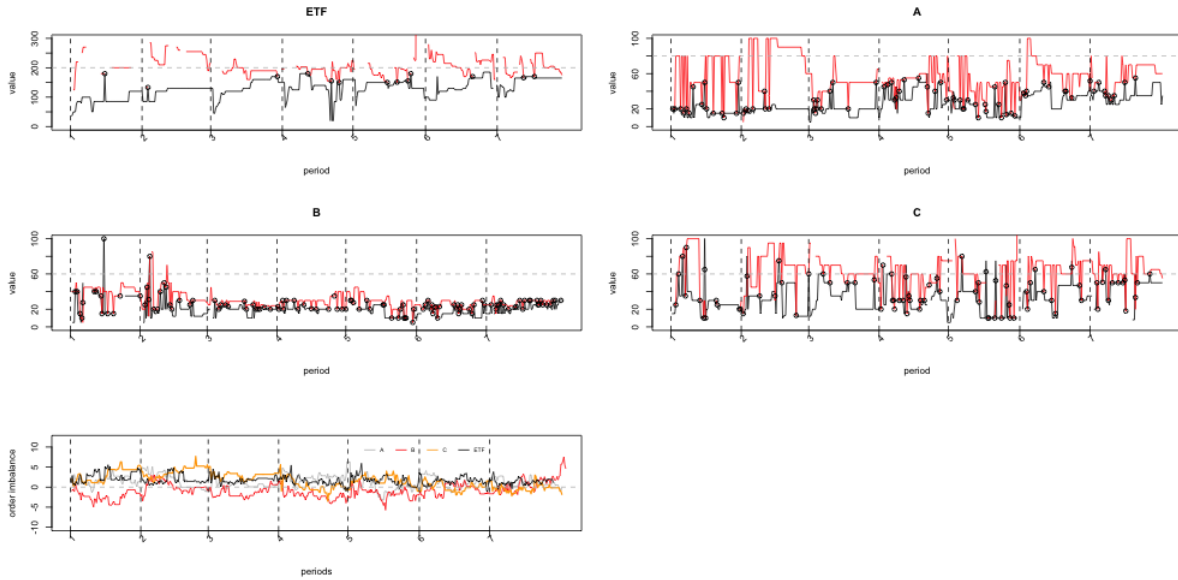


Figure 17: Best bids/offers, prices and order imbalance for $A2C_{short}$ treatment (session 3)

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).

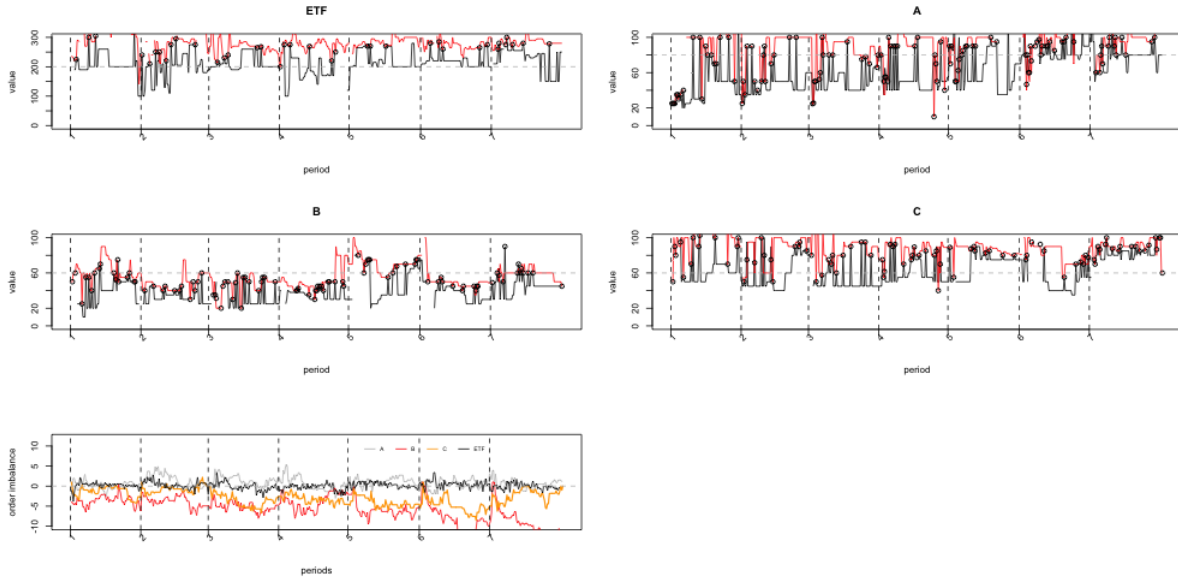


Figure 18: Best bids/offers, prices and order imbalance for $A2C_{short}$ treatment (session 4)

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).

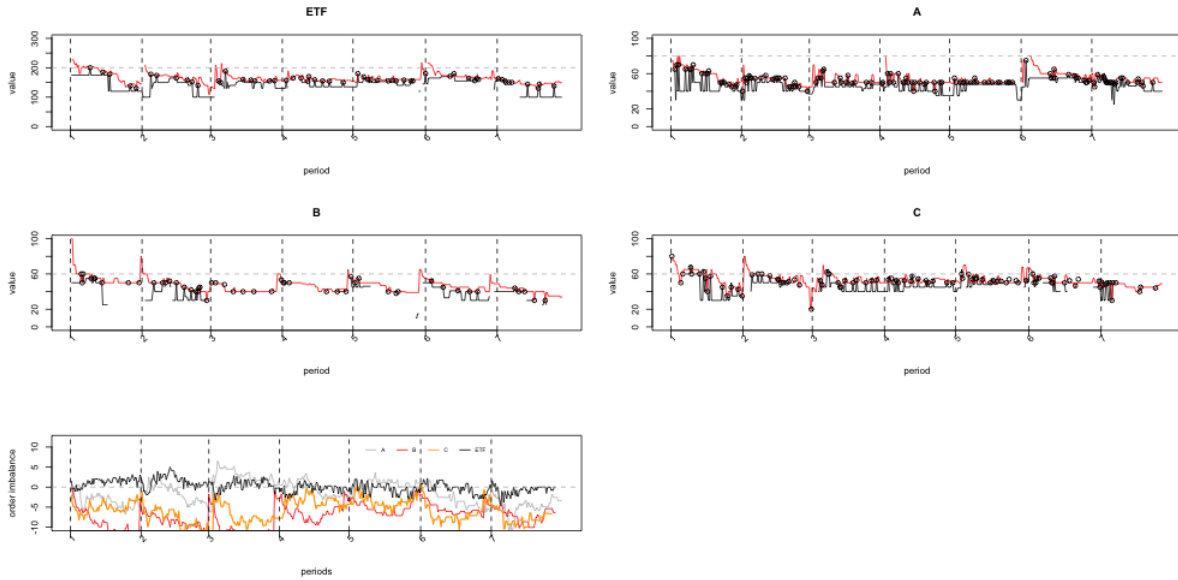


Figure 19: Best bids/offers, prices and order imbalance for $A2C_{short}$ treatment (session 5)

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).