ETF Short Interest and Failures-to-Deliver: Naked Short-Selling or Operational Shorting?

Richard B. Evans Rabih Moussawi Michael S. Pagano John Sedunov

Discussant Shaun Davies – University of Colorado, Boulder



 Understanding the arbitrage mechanism in ETFs is important for participants, regulators, and academics



- Understanding the arbitrage mechanism in ETFs is important for participants, regulators, and academics
- Concept of "operational shorting" is intuitive and compelling



- Understanding the arbitrage mechanism in ETFs is important for participants, regulators, and academics
- Concept of "operational shorting" is intuitive and compelling
 - A basic understanding of option value suggests it is almost certainly a common practice



- Understanding the arbitrage mechanism in ETFs is important for participants, regulators, and academics
- Concept of "operational shorting" is intuitive and compelling
 - A basic understanding of option value suggests it is almost certainly a common practice
- My discussion is twofold:



- Understanding the arbitrage mechanism in ETFs is important for participants, regulators, and academics
- Concept of "operational shorting" is intuitive and compelling
 - A basic understanding of option value suggests it is almost certainly a common practice
- My discussion is twofold:
 - (1) **For the Authors**: Provide a parsimonious model of the ETF arbitrage mechanism to discipline the empirical work and provide new testable predictions



- Understanding the arbitrage mechanism in ETFs is important for participants, regulators, and academics
- Concept of "operational shorting" is intuitive and compelling
 - A basic understanding of option value suggests it is almost certainly a common practice
- My discussion is twofold:
 - (1) **For the Authors**: Provide a parsimonious model of the ETF arbitrage mechanism to discipline the empirical work and provide new testable predictions
 - (2) **For the Audience**: Build the model from pseudo first principles to build greater appreciation for ETFs in general



- Understanding the arbitrage mechanism in ETFs is important for participants, regulators, and academics
- Concept of "operational shorting" is intuitive and compelling
 - A basic understanding of option value suggests it is almost certainly a common practice
- My discussion is twofold:
 - (1) For the Authors: Provide a parsimonious model of the ETF arbitrage mechanism to discipline the empirical work and provide new testable predictions
 - (2) For the Audience: Build the model from pseudo first principles to build greater appreciation for ETFs in general
- My Big Point: Operational shorting substitutes for pure AP activity \Rightarrow Can authors examine the extent? What about relative price efficiency?

CFIC 2018



 ETFs are designed to maintain relative price efficiency between the ETF shares and the underlying assets



- ETFs are designed to maintain relative price efficiency between the ETF shares and the underlying assets
- The creation/redemption process enables arbitrageurs to earn arbitrage profits...



- ETFs are designed to maintain relative price efficiency between the ETF shares and the underlying assets
- The creation/redemption process enables arbitrageurs to earn arbitrage profits...
- ... the arbitrage profits are a means to restore relative price efficiency between the ETF and the underlying



- ETFs are designed to maintain relative price efficiency between the ETF shares and the underlying assets
- The creation/redemption process enables arbitrageurs to earn arbitrage profits...
- ... the arbitrage profits are a means to restore relative price efficiency between the ETF and the underlying
- Importantly, this implies that relative price efficiency is restored by affecting the supply of shares outstanding



- ETFs are designed to maintain relative price efficiency between the ETF shares and the underlying assets
- The creation/redemption process enables arbitrageurs to earn arbitrage profits...
- ... the arbitrage profits are a means to restore relative price efficiency between the ETF and the underlying
- Importantly, this implies that relative price efficiency is restored by affecting the supply of shares outstanding
- Implicitly, the ETF mechanism assumes that ETF investors' demand is downward sloping in the short-run



- ETFs are designed to maintain relative price efficiency between the ETF shares and the underlying assets
- The creation/redemption process enables arbitrageurs to earn arbitrage profits...
- ... the arbitrage profits are a means to restore relative price efficiency between the ETF and the underlying
- Importantly, this implies that relative price efficiency is restored by affecting the supply of shares outstanding
- Implicitly, the ETF mechanism assumes that ETF investors' demand is downward sloping in the short-run and/or demand for the underlying is downward sloping in the short-run



• Two period model



CFIC 2018

- Two period model
- A passively managed equity ETF managed by a risk neutral, competitive provider (e.g., BlackRock)



- Two period model
- A passively managed equity ETF managed by a risk neutral, competitive provider (e.g., BlackRock)
- N Authorized Participants (APs)



- Two period model
- A passively managed equity ETF managed by a risk neutral, competitive provider (e.g., BlackRock)
- N Authorized Participants (APs)
 - APs can exploit mispricing via creating/redeeming shares



- Two period model
- A passively managed equity ETF managed by a risk neutral, competitive provider (e.g., BlackRock)
- N Authorized Participants (APs)
 - APs can exploit mispricing via creating/redeeming shares
 - APs can exploit mispricing via operational shorting
 - APs incur quadratic risk cost (e.g., unhedged price movements)



- Two period model
- A passively managed equity ETF managed by a risk neutral, competitive provider (e.g., BlackRock)
- N Authorized Participants (APs)
 - APs can exploit mispricing via creating/redeeming shares
 - APs can exploit mispricing via operational shorting
 - APs incur quadratic risk cost (e.g., unhedged price movements)
- The ETF mimics an underlying asset χ (e.g., the S&P 500 Index)



- Two period model
- A passively managed equity ETF managed by a risk neutral, competitive provider (e.g., BlackRock)
- N Authorized Participants (APs)
 - APs can exploit mispricing via creating/redeeming shares
 - APs can exploit mispricing via operational shorting
 - APs incur quadratic risk cost (e.g., unhedged price movements)
- The ETF mimics an underlying asset χ (e.g., the S&P 500 Index)
- The initial number of ETF shares is of measure length q₀



- Two period model
- A passively managed equity ETF managed by a risk neutral, competitive provider (e.g., BlackRock)
- N Authorized Participants (APs)
 - APs can exploit mispricing via creating/redeeming shares
 - APs can exploit mispricing via operational shorting
 - APs incur quadratic risk cost (e.g., unhedged price movements)
- The ETF mimics an underlying asset χ (e.g., the S&P 500 Index)
- The initial number of ETF shares is of measure length q₀
- The ETF shares' price *p_t* and underlying asset's tradable value (i.e., NAV) π_t are endogenously determined



- Two period model
- A passively managed equity ETF managed by a risk neutral, competitive provider (e.g., BlackRock)
- N Authorized Participants (APs)
 - APs can exploit mispricing via creating/redeeming shares
 - APs can exploit mispricing via operational shorting
 - APs incur quadratic risk cost (e.g., unhedged price movements)
- The ETF mimics an underlying asset χ (e.g., the S&P 500 Index)
- The initial number of ETF shares is of measure length q₀
- The ETF shares' price *p_t* and underlying asset's tradable value (i.e., NAV) π_t are endogenously determined

•
$$\psi_t \equiv p_t - \pi_t \Rightarrow \mathsf{ETF}$$
 premium



Model Timing

(1) The ETF and underlying asset are efficiently priced at t = 0

- (a) Demand shock hits both ETF and underlying assets (but to different degrees),
- (b) APs step in and exploit mispricing
 - (i) Create (redeem) shares to exploit arbitrage,
 - (ii) and/or operational shorting to exploit arbitrage
- (c) AP arbitrage activity affects price levels of both ETF and underlying asset,
- (2) The ETF's price and underlying asset's NAV are established at t = 1,
 - (a) The ETF premium is finalized



• ETF investors' collective demand is a downward-sloped curve,

$$\boldsymbol{p}_t = -\phi \boldsymbol{q}_t + \boldsymbol{\epsilon}_t + \boldsymbol{\alpha},$$



• ETF investors' collective demand is a downward-sloped curve,

$$\boldsymbol{p}_t = -\phi \boldsymbol{q}_t + \boldsymbol{\epsilon}_t + \boldsymbol{\alpha},$$

• $\phi \in \mathbb{R}^+$ is the sensitivity of the ETF share price to changes in ETF share quantity



• ETF investors' collective demand is a downward-sloped curve,

$$\boldsymbol{p}_t = -\phi \boldsymbol{q}_t + \boldsymbol{\epsilon}_t + \boldsymbol{\alpha},$$

- $\phi \in \mathbb{R}^+$ is the sensitivity of the ETF share price to changes in ETF share quantity
 - Lower values of φ relate to better liquidity in the ETF shares, e.g., larger investor base and less price impact from share creation/redemption.



• ETF investors' collective demand is a downward-sloped curve,

$$\boldsymbol{p}_t = -\phi \boldsymbol{q}_t + \boldsymbol{\epsilon}_t + \boldsymbol{\alpha},$$

- $\phi \in \mathbb{R}^+$ is the sensitivity of the ETF share price to changes in ETF share quantity
 - Lower values of φ relate to better liquidity in the ETF shares, e.g., larger investor base and less price impact from share creation/redemption.
- ϵ_t is a demand shock ($\epsilon_0 = 0$)



• ETF investors' collective demand is a downward-sloped curve,

 $\boldsymbol{p}_t = -\phi \boldsymbol{q}_t + \boldsymbol{\epsilon}_t + \boldsymbol{\alpha},$

- $\phi \in \mathbb{R}^+$ is the sensitivity of the ETF share price to changes in ETF share quantity
 - Lower values of φ relate to better liquidity in the ETF shares, e.g., larger investor base and less price impact from share creation/redemption.
- ϵ_t is a demand shock ($\epsilon_0 = 0$)
- α is an arbitrary constant that ensures initial share quantity is strictly positive



• ETF investors' collective demand is a downward-sloped curve,

 $\boldsymbol{p}_t = -\phi \boldsymbol{q}_t + \boldsymbol{\epsilon}_t + \boldsymbol{\alpha},$

- $\phi \in \mathbb{R}^+$ is the sensitivity of the ETF share price to changes in ETF share quantity
 - Lower values of φ relate to better liquidity in the ETF shares, e.g., larger investor base and less price impact from share creation/redemption.
- ϵ_t is a demand shock ($\epsilon_0 = 0$)
- α is an arbitrary constant that ensures initial share quantity is strictly positive
- Underlying asset NAV is π_0 and $p_0 = \pi_0$



• ETF investors' collective demand is a downward-sloped curve,

 $\boldsymbol{p}_t = -\phi \boldsymbol{q}_t + \boldsymbol{\epsilon}_t + \boldsymbol{\alpha},$

- $\phi \in \mathbb{R}^+$ is the sensitivity of the ETF share price to changes in ETF share quantity
 - Lower values of φ relate to better liquidity in the ETF shares, e.g., larger investor base and less price impact from share creation/redemption.
- ϵ_t is a demand shock ($\epsilon_0 = 0$)
- α is an arbitrary constant that ensures initial share quantity is strictly positive
- Underlying asset NAV is π_0 and $p_0 = \pi_0$

• t = 1: ϵ_1 is drawn from $g(\epsilon_1)$ on the support $(-\infty, \infty)$



The APs collectively affect demand for shares and underlying



- The APs collectively affect demand for shares and underlying
 - Each AP *i* chooses δ_i^{AP} shares to create/redeem and δ_i^{OS} to operationally short



- The APs collectively affect demand for shares and underlying
 - Each AP i chooses δ^{AP}_i shares to create/redeem and δ^{OS}_i to operationally short
 - Total quantity to affect ETF demand is

$$\Delta_t^{AP} + \Delta_t^{OS} \equiv \sum_{i=1}^{N} \left(\delta_i^{AP} + \delta_i^{OS} \right)$$

• Total quantity to affect underlying demand is $\Delta_t^{AP} \equiv \sum_{i=1}^N \delta_i^{AP}$



- The APs collectively affect demand for shares and underlying
 - Each AP i chooses δ^{AP}_i shares to create/redeem and δ^{OS}_i to operationally short
 - Total quantity to affect ETF demand is

$$\Delta_t^{AP} + \Delta_t^{OS} \equiv \sum_{i=1}^N \left(\delta_i^{AP} + \delta_i^{OS} \right)$$

- Total quantity to affect underlying demand is $\Delta_t^{AP} \equiv \sum_{i=1}^N \delta_i^{AP}$
- The tradable NAV price π_t is a function of both ϵ_t and Δ_t^{AP} ,

$$\pi_t = \pi_{t-1} + \beta \epsilon_t + \lambda \Delta_t^{AP}.$$



- The APs collectively affect demand for shares and underlying
 - Each AP i chooses δ^{AP}_i shares to create/redeem and δ^{OS}_i to operationally short
 - Total quantity to affect ETF demand is

$$\Delta_t^{AP} + \Delta_t^{OS} \equiv \sum_{i=1}^N \left(\delta_i^{AP} + \delta_i^{OS} \right)$$

- Total quantity to affect underlying demand is $\Delta_t^{AP} \equiv \sum_{i=1}^N \delta_i^{AP}$
- The tradable NAV price π_t is a function of both ϵ_t and Δ_t^{AP} ,

$$\pi_t = \pi_{t-1} + \beta \epsilon_t + \lambda \Delta_t^{AP}.$$

β ∈ [0, 1] allows demand shocks to also affect the underlying assets



- The APs collectively affect demand for shares and underlying
 - Each AP i chooses δ^{AP}_i shares to create/redeem and δ^{OS}_i to operationally short
 - Total quantity to affect ETF demand is

$$\Delta_t^{AP} + \Delta_t^{OS} \equiv \sum_{i=1}^N \left(\delta_i^{AP} + \delta_i^{OS} \right)$$

- Total quantity to affect underlying demand is $\Delta_t^{AP} \equiv \sum_{i=1}^{N} \delta_i^{AP}$
- The tradable NAV price π_t is a function of both ϵ_t and Δ_t^{AP} ,

$$\pi_t = \pi_{t-1} + \beta \epsilon_t + \lambda \Delta_t^{AP}.$$

- $\beta \in [0, 1]$ allows demand shocks to also affect the underlying assets
- λ relates buying (or selling) of underlying by arbitrageurs into price impact

- The APs collectively affect demand for shares and underlying
 - Each AP i chooses δ^{AP}_i shares to create/redeem and δ^{OS}_i to operationally short
 - Total quantity to affect ETF demand is

$$\Delta_t^{AP} + \Delta_t^{OS} \equiv \sum_{i=1}^N \left(\delta_i^{AP} + \delta_i^{OS} \right)$$

- Total quantity to affect underlying demand is $\Delta_t^{AP} \equiv \sum_{i=1}^N \delta_i^{AP}$
- The tradable NAV price π_t is a function of both ϵ_t and Δ_t^{AP} ,

$$\pi_t = \pi_{t-1} + \beta \epsilon_t + \lambda \Delta_t^{AP}.$$

- β ∈ [0, 1] allows demand shocks to also affect the underlying assets
- λ relates buying (or selling) of underlying by arbitrageurs into price impact
 - For example, linear pricing rule in Kyle (1985)

Arbitrage Activity

• Each AP's optimal creation/redemption choice solves,

$$\max_{\substack{\delta_{i}^{AP}, \delta_{i}^{OS} \in \mathbb{R} \\ 0}} \underbrace{\delta_{i}^{AP} \left(p_{1} \left(\delta_{i}^{AP} + \delta_{i}^{OS} + \delta_{-i}^{AP} + \delta_{-i}^{OS} \right) - \pi_{1} \left(\delta_{i}^{AP} + + \delta_{-i}^{AP} \right) \right)}_{\text{AP Activity Profits}} + \underbrace{\delta_{i}^{OS} \left(p_{1} \left(\delta_{i}^{AP} + \delta_{i}^{OS} + \delta_{-i}^{AP} + \delta_{-i}^{OS} \right) - \pi_{1} \left(\delta_{i}^{AP} + + \delta_{-i}^{AP} \right) \right)}_{\text{Operational Shorting Profits}} - \underbrace{\omega \frac{\delta_{i}^{OS^{2}}}{2}}_{\text{Operational Shorting Risk}}$$



Arbitrage Activity

• Each AP's optimal creation/redemption choice solves,

$$\max_{\substack{\delta_{i}^{AP}, \delta_{i}^{OS} \in \mathbb{R} \\ 0}} \underbrace{\delta_{i}^{AP} \left(p_{1}(\delta_{i}^{AP} + \delta_{i}^{OS} + \delta_{-i}^{AP} + \delta_{-i}^{OS}) - \pi_{1}(\delta_{i}^{AP} + + \delta_{-i}^{AP}) \right)}_{\text{AP Activity Profits}}$$

$$+ \underbrace{\delta_{i}^{OS} \left(p_{1}(\delta_{i}^{AP} + \delta_{i}^{OS} + \delta_{-i}^{AP} + \delta_{-i}^{OS}) - \pi_{1}(\delta_{i}^{AP} + + \delta_{-i}^{AP}) \right)}_{\text{Operational Shorting Profits}}$$

$$- \underbrace{\omega_{i}^{\frac{\delta_{i}^{OS}^{2}}{2}}}_{\text{Operational Shorting Risk}}$$

 AP's choice partially internalizes effects on both the ETF and underlying asset prices

CFIC 2018

Equilibrium AP Trades

• Assume $\omega \geq \lambda$ for tractability



CFIC 2018

Equilibrium AP Trades

- Assume $\omega \geq \lambda$ for tractability
- The equilibrium **AP-level** redemption/creation and **aggregate** redemption/creation activity are,

$$\delta_{i}^{AP} = \frac{(1-\beta)(\omega-\lambda)\epsilon}{(N+1)\omega(\lambda+\phi) - N\lambda^{2}}, \ \Delta^{AP} = \frac{N(1-\beta)(\omega-\lambda)\epsilon}{(N+1)\omega(\lambda+\phi) - N\lambda^{2}}$$
$$\delta_{i}^{OS} = \frac{(1-\beta)\lambda\epsilon}{(N+1)\omega(\lambda+\phi) - N\lambda^{2}}, \ \Delta^{OS} = \frac{N(1-\beta)\lambda\epsilon}{(N+1)\omega(\lambda+\phi) - N\lambda^{2}}$$



Equilibrium AP Trades

- Assume $\omega \geq \lambda$ for tractability
- The equilibrium **AP-level** redemption/creation and **aggregate** redemption/creation activity are,

$$\delta_{i}^{AP} = \frac{(1-\beta)(\omega-\lambda)\epsilon}{(N+1)\omega(\lambda+\phi) - N\lambda^{2}}, \ \Delta^{AP} = \frac{N(1-\beta)(\omega-\lambda)\epsilon}{(N+1)\omega(\lambda+\phi) - N\lambda^{2}}$$
$$\delta_{i}^{OS} = \frac{(1-\beta)\lambda\epsilon}{(N+1)\omega(\lambda+\phi) - N\lambda^{2}}, \ \Delta^{OS} = \frac{N(1-\beta)\lambda\epsilon}{(N+1)\omega(\lambda+\phi) - N\lambda^{2}}$$

• Equilibrium premium is given by:

$$\psi_t \equiv \rho_t - \pi_t = \frac{(1-\beta)\omega(\lambda+\phi)\epsilon}{(N+1)\omega(\lambda+\phi) - N\lambda^2}$$



Comparative Statics of AP and OS Activity (1) $\uparrow \lambda \Rightarrow \uparrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$



CFIC 2018

Page 10

- (1) $\uparrow \lambda \Rightarrow \uparrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$
 - More price impact in underlying, more operational shorting and less AP activity
 - EP: Less liquid underlying should be characterized by more operational shorting and less AP activity (substitution effect)



- (1) $\uparrow \lambda \Rightarrow \uparrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$
 - More price impact in underlying, more operational shorting and less AP activity
 - EP: Less liquid underlying should be characterized by more operational shorting and less AP activity (substitution effect)

(2) $\uparrow \omega \Rightarrow \downarrow \Delta^{OS}$ and $\uparrow \Delta^{AP}$



- (1) $\uparrow \lambda \Rightarrow \uparrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$
 - More price impact in underlying, more operational shorting and less AP activity
 - **EP**: Less liquid underlying should be characterized by more operational shorting and less AP activity (substitution effect)

(2) $\uparrow \omega \Rightarrow \downarrow \Delta^{OS}$ and $\uparrow \Delta^{AP}$

- Larger cost to operational shorting, less operational shorting and more AP activity
- **EP**: Hard to hedge underlying should be characterized by less operational shorting and more AP activity (substitution effect)



- (1) $\uparrow \lambda \Rightarrow \uparrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$
 - More price impact in underlying, more operational shorting and less AP activity
 - **EP**: Less liquid underlying should be characterized by more operational shorting and less AP activity (substitution effect)

(2) $\uparrow \omega \Rightarrow \downarrow \Delta^{OS}$ and $\uparrow \Delta^{AP}$

- Larger cost to operational shorting, less operational shorting and more AP activity
- **EP**: Hard to hedge underlying should be characterized by less operational shorting and more AP activity (substitution effect)

(3) $\uparrow \phi \Rightarrow \downarrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$



- (1) $\uparrow \lambda \Rightarrow \uparrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$
 - More price impact in underlying, more operational shorting and less AP activity
 - EP: Less liquid underlying should be characterized by more operational shorting and less AP activity (substitution effect)

(2) $\uparrow \omega \Rightarrow \downarrow \Delta^{OS}$ and $\uparrow \Delta^{AP}$

- Larger cost to operational shorting, less operational shorting and more AP activity
- EP: Hard to hedge underlying should be characterized by less operational shorting and more AP activity (substitution effect)

(3) $\uparrow \phi \Rightarrow \downarrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$

- More inelastic demand for ETF shares, less operational shorting and less AP activity
- EP: Liquid ETFs with diverse clienteles should have more operational shorting and more AP activity

(1) $\uparrow \beta \Rightarrow \downarrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$



- (1) $\uparrow \beta \Rightarrow \downarrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$
 - More similar clienteles for ETF shares and underlying, less operational shorting and less AP activity
 - EP: More similar ownership, e.g., institutional ownership of underlying relative to institutional ownership of ETF shares, should have relatively less operational shorting and less AP activity



- (1) $\uparrow \beta \Rightarrow \downarrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$
 - More similar clienteles for ETF shares and underlying, less operational shorting and less AP activity
 - EP: More similar ownership, e.g., institutional ownership of underlying relative to institutional ownership of ETF shares, should have relatively less operational shorting and less AP activity

(2)
$$\uparrow N \Rightarrow \uparrow \Delta^{OS}$$
 and $\uparrow \Delta^{AP}$



- (1) $\uparrow \beta \Rightarrow \downarrow \Delta^{OS}$ and $\downarrow \Delta^{AP}$
 - More similar clienteles for ETF shares and underlying, less operational shorting and less AP activity
 - EP: More similar ownership, e.g., institutional ownership of underlying relative to institutional ownership of ETF shares, should have relatively less operational shorting and less AP activity

(2)
$$\uparrow N \Rightarrow \uparrow \Delta^{OS}$$
 and $\uparrow \Delta^{AP}$

- More APs, internalize a smaller fraction of AP and OS activity and trade more
- **EP**: ETFs with more APs should have relatively more operational shorting and more AP activity
- Model is generally consistent with paper!



(1) $\uparrow \lambda \Rightarrow \uparrow \psi_t$



(1) $\uparrow \lambda \Rightarrow \uparrow \psi_t$

(a) More price impact in underlying, AP activity is substituted for, but not fully



- (1) $\uparrow \lambda \Rightarrow \uparrow \psi_t$
 - (a) More price impact in underlying, AP activity is substituted for, but not fully
- (2) $\uparrow \omega \Rightarrow \downarrow \psi_t$



- (1) $\uparrow \lambda \Rightarrow \uparrow \psi_t$
 - (a) More price impact in underlying, AP activity is substituted for, but not fully
- (2) $\uparrow \omega \Rightarrow \downarrow \psi_t$
 - (a) Larger cost to operational shorting, greater resulting AP activity is more effective in restoring relative price efficiency
 - (b) Price impact on **both** the underlying and the ETF shares



- (1) $\uparrow \lambda \Rightarrow \uparrow \psi_t$
 - (a) More price impact in underlying, AP activity is substituted for, but not fully
- (2) $\uparrow \omega \Rightarrow \downarrow \psi_t$
 - (a) Larger cost to operational shorting, greater resulting AP activity is more effective in restoring relative price efficiency
 - (b) Price impact on **both** the underlying and the ETF shares
- $(3) \uparrow \phi \Rightarrow \downarrow \psi_t$



- (1) $\uparrow \lambda \Rightarrow \uparrow \psi_t$
 - (a) More price impact in underlying, AP activity is substituted for, but not fully
- (2) $\uparrow \omega \Rightarrow \downarrow \psi_t$
 - (a) Larger cost to operational shorting, greater resulting AP activity is more effective in restoring relative price efficiency
 - (b) Price impact on **both** the underlying and the ETF shares
- $(3) \uparrow \phi \Rightarrow \downarrow \psi_t$
 - (a) More inelastic demand for ETF shares ⇒ bigger price impact on ETF shares, better relative price efficiency



- (1) $\uparrow \lambda \Rightarrow \uparrow \psi_t$
 - (a) More price impact in underlying, AP activity is substituted for, but not fully
- (2) $\uparrow \omega \Rightarrow \downarrow \psi_t$
 - (a) Larger cost to operational shorting, greater resulting AP activity is more effective in restoring relative price efficiency
 - (b) Price impact on **both** the underlying and the ETF shares
- $(3) \uparrow \phi \Rightarrow \downarrow \psi_t$
 - (a) More inelastic demand for ETF shares ⇒ bigger price impact on ETF shares, better relative price efficiency

 $(4) \uparrow \beta \Rightarrow \downarrow \psi_t$



- (1) $\uparrow \lambda \Rightarrow \uparrow \psi_t$
 - (a) More price impact in underlying, AP activity is substituted for, but not fully
- (2) $\uparrow \omega \Rightarrow \downarrow \psi_t$
 - (a) Larger cost to operational shorting, greater resulting AP activity is more effective in restoring relative price efficiency
 - (b) Price impact on **both** the underlying and the ETF shares
- $(3) \uparrow \phi \Rightarrow \downarrow \psi_t$
 - (a) More inelastic demand for ETF shares ⇒ bigger price impact on ETF shares, better relative price efficiency
- $(4) \uparrow \beta \Rightarrow \downarrow \psi_t$
 - (a) More similar clienteles, better relative price efficiency



- (1) $\uparrow \lambda \Rightarrow \uparrow \psi_t$
 - (a) More price impact in underlying, AP activity is substituted for, but not fully
- (2) $\uparrow \omega \Rightarrow \downarrow \psi_t$
 - (a) Larger cost to operational shorting, greater resulting AP activity is more effective in restoring relative price efficiency
 - (b) Price impact on **both** the underlying and the ETF shares
- $(3) \uparrow \phi \Rightarrow \downarrow \psi_t$
 - (a) More inelastic demand for ETF shares ⇒ bigger price impact on ETF shares, better relative price efficiency
- $(4) \uparrow \beta \Rightarrow \downarrow \psi_t$
 - (a) More similar clienteles, better relative price efficiency
- (5) $\uparrow \mathbf{N} \Rightarrow \downarrow \psi_t$



- (1) $\uparrow \lambda \Rightarrow \uparrow \psi_t$
 - (a) More price impact in underlying, AP activity is substituted for, but not fully
- (2) $\uparrow \omega \Rightarrow \downarrow \psi_t$
 - (a) Larger cost to operational shorting, greater resulting AP activity is more effective in restoring relative price efficiency
 - (b) Price impact on **both** the underlying and the ETF shares
- $(3) \uparrow \phi \Rightarrow \downarrow \psi_t$
 - (a) More inelastic demand for ETF shares ⇒ bigger price impact on ETF shares, better relative price efficiency
- $(4) \uparrow \beta \Rightarrow \downarrow \psi_t$
 - (a) More similar clienteles, better relative price efficiency
- (5) $\uparrow \mathbf{N} \Rightarrow \downarrow \psi_t$
 - (a) More APs, better relative price efficiency



Operational shorting substitutes for traditional AP creation/redemption activity



- Operational shorting substitutes for traditional AP creation/redemption activity
- Operational shorting is not as effective in restoring price efficiency



- Operational shorting substitutes for traditional AP creation/redemption activity
- Operational shorting is not as effective in restoring price efficiency
 - Creation/redemption activity puts price pressure on both the underlying assets and the ETF shares



- Operational shorting substitutes for traditional AP creation/redemption activity
- Operational shorting is not as effective in restoring price efficiency
 - Creation/redemption activity puts price pressure on both the underlying assets and the ETF shares
 - Caveat: requires that price impact on underlying assets via hedging is **not equal** to outright purchase of underlying



- Operational shorting substitutes for traditional AP creation/redemption activity
- Operational shorting is not as effective in restoring price efficiency
 - Creation/redemption activity puts price pressure on both the underlying assets and the ETF shares
 - Caveat: requires that price impact on underlying assets via hedging is **not equal** to outright purchase of underlying
 - Derivative-based ETFs' price efficiency may be the same with operational shorting versus without (because the underlying asset is the same/or similar to the hedging instrument)



- Operational shorting substitutes for traditional AP creation/redemption activity
- Operational shorting is not as effective in restoring price efficiency
 - Creation/redemption activity puts price pressure on both the underlying assets and the ETF shares
 - Caveat: requires that price impact on underlying assets via hedging is **not equal** to outright purchase of underlying
 - Derivative-based ETFs' price efficiency may be the same with operational shorting versus without (because the underlying asset is the same/or similar to the hedging instrument)
- Authors may find interesting results looking at the composition of arbitrage activity (fraction that is operational shorting and fraction that is traditional AP creation/redemption activity)

