

Supplemental Appendix to “A Long and a Short Leg Make For a Wobbly Equilibrium”

Nicolae Gârleanu¹, Stavros Panageas², and Geoffery Zheng³

¹WashU, Olin Bsuiness School and NBER

²UCLA, Anderson School of Management and NBER

³NYU Shanghai

A The Determination of the Lending Fee

In the text we assume a “flat” supply curve for lending shares. That is, we assume $f_t = f(y_t) = \varphi$. We provide here the simplest model that supports this assumption. We also discuss how to extend the model to allow for an increasing $f(\cdot)$.

All interactions considered in this section happen anew every period, where the length of the period is idealized to be “ dt ,” that is, infinitesimal. (We could formalize this assumption by considering a discrete-time model where the length Δ of a period is taken to go to zero, and focusing on the limit of resultant equilibria.)

We start by considering the long investors, who wish to lend their shares. Each investor lends all her shares to any one of a competitive fringe of profit-maximizing “security lenders” in exchange for an income stream that is proportional to the dollar value of shares the investor lends. This income stream is determined as follows. In equilibrium, each security lender lends a proportion y_t of the shares it borrows from investors and receives a fee f_l per dollar of shares it lends out. (We omit time subscripts from now on.) Competition between the security lenders drives the income stream paid to long investors to $y f_l$ per dollar of stock owned by the investors.⁵⁵

At the other end of the lending transaction, desirous short sellers interact with a competitive set of “borrower’s brokers.” Specifically, for every borrowing fee f_b the would-be short sellers provide the dollar amount that they would like to short, and the brokers take the value f_b as a given when they attempt to fill the investors’ borrowing orders.

All of the frictions in this model pertain to the interaction between security lenders and borrower’s brokers. Specifically, to initiate a stock loan the representative broker must pay a cost ξ per dollar value of share “located” with a security lender, per unit of time. This cost is construed as labor cost that compensates brokers for their disutility of labor.

The interaction between the broker and the security lender takes the form of bilateral Nash bargaining in which the broker has bargaining power $1/(1+z)$ for a parameter $z \in (0, \infty)$. Given our assumption that all interactions (between investors and brokers or security lenders and between brokers and security lenders) happen anew every period, the outside option for both brokers and security lenders is the failure to transact during the period. This means

55. In that sense, the security lenders resemble the “insurance companies” in Blanchard (1985). Similar to how insurance companies collect payments from the fraction of agents who die and rebate them to the surviving population, the security lenders collect lending fees from the proportion of a long portfolio that gets loaned out and rebate it in the form of an income stream to the representative long investor.

that the gains from trade to the security lender equal the lending fee f_l , while to the broker the borrowing fee net of the lending one $f_b - f_l$ — the searching and matching cost ξ has been sunk at this point. The total gains from trade equal f_b , the foregone revenue from the would-be short seller. Given the bargaining protocol, it follows that

$$f_l = \frac{z}{1+z}(f_l + f_b - f_l) = \frac{z}{1+z}f_b. \quad (\text{A.1})$$

Brokers break even on net, meaning that

$$f_b = f_l + \xi, \quad (\text{A.2})$$

so that

$$f_l = z\xi, \quad (\text{A.3})$$

$$f_b = (1+z)\xi. \quad (\text{A.4})$$

To keep the model transparent and tractable, assume that all brokers are members of the representative household, and therefore the fees that compensate them for their effort are rebated to each households as an income stream proportional to the household's wealth and independent of the composition of the household's portfolio.

Setting $\varphi = (1+z)\xi$ and $\tau = z/(1+z)$, this extended model is equivalent to the model we assumed in the text. To generalize to upward-sloping supply curves, one would simply assume an increasing cost $\xi(y)$.

B Multiple Agent Types

We illustrate here that the multiplicity of equilibria may expand with the number of agent types. In particular, adding a third group of agents can result in a third equilibrium featuring non-zero shorting; such a model may admit, in fact, up to five equilibria.

Specifically, let us assume a third group of investors characterized by beliefs that are summarized by the quantity η^P . We think of these investors as pessimists, which implies $\eta^P < 0$. The intuition we wish to capture is that, in addition to the “high-shorting” and “medium-shorting” equilibria in the base-line model, low-shorting equilibria may exist in which investor R is inactive, while investor P shorts actively.

To make the point theoretically, one may argue by continuity. Specifically, consider the zero-shorting equilibrium in the baseline model, and perturb the setting by adding a small mass of sufficiently pessimistic investors ($|\eta^P|$ large enough). These investors will want to short, but will not be sufficiently numerous to move the Sharpe ratio or lending income to a point where investors R and I are no longer in equilibrium.

It is helpful to write down the equilibrium conditions in the augmented model — both to allow for a formal argument and in the interest of a numerical illustration. We repeat the analysis in the text — letting ω^P denote the wealth share of agents P — to obtain the

market clearing condition

$$1 = \frac{1}{\sigma_D} \left[\omega^P \left(\kappa + \eta^P + \frac{\varphi}{\sigma_D} \right) 1_{\{\kappa + \eta^P + \frac{\varphi}{\sigma_D} < 0\}} + \omega^R \left(\kappa + \frac{\varphi}{\sigma_D} \right) 1_{\{\kappa + \frac{\varphi}{\sigma_D} < 0\}} + \omega^I \left(\kappa + \eta^I + \frac{\varphi}{\sigma_D} \tau y \right) \right], \quad (\text{B.1})$$

where the left-hand side is the proportion of aggregate wealth represented by the supply of the stock, while the right-hand side equals the proportion of aggregate wealth invested in the stock. We restricted attention to cases in which R agents do not take a long position in the stock.

We solve for the Sharpe ratio κ :

$$\kappa = \sigma_D - (\omega^P \eta^P + \omega^I \eta^I) - \frac{\varphi}{\sigma_D} (\omega^P + \omega^R + \omega^I \tau y) \quad (\text{B.2})$$

if $\kappa + \varphi/\sigma_D < 0$, respectively

$$\kappa = \frac{\sigma_D}{\omega^P + \omega^I} - \frac{\omega^P \eta^P + \omega^I \eta^I}{\omega^P + \omega^I} - \frac{\varphi}{\sigma_D} \frac{\omega^P + \omega^I \tau y}{\omega^P + \omega^I} \quad (\text{B.3})$$

if $\kappa + \varphi/\sigma_D \geq 0 > \kappa + \eta^P + \varphi/\sigma_D$.

The other equilibrium condition concerns the determination of the value of y :

$$y = - \frac{\omega^P \left(\kappa + \eta^P + \frac{\varphi}{\sigma_D} \right) 1_{\{\kappa + \eta^P + \frac{\varphi}{\sigma_D} < 0\}} + \omega^R \left(\kappa + \frac{\varphi}{\sigma_D} \right) 1_{\{\kappa + \frac{\varphi}{\sigma_D} < 0\}}}{\omega^I \left(\kappa + \eta^I + \frac{\varphi}{\sigma_D} \tau y \right)}. \quad (\text{B.4})$$

Depending on whether κ is determined according to (B.2) or (B.3) we obtain a different quadratic equation. For appropriate parameter choices all but one combinations are possible in terms of how many solutions in the interval $(0, 1)$ each of them admits. We are particularly interested in situations in which (B.2) applies and results in two admissible solutions, in addition to which at least one solution obtains when (B.3) applies.

We illustrate such outcomes in Figure B.1. The two panels differ in terms of parameters, but depict the same objects. Specifically, the x-axis records candidate values of y that agents anticipate. Agents form demands taking such a value y and a Sharpe ratio κ as given, and clearing in the asset market determines the Sharpe ratio. With the Sharpe ratio now specified for each candidate y , we can compute the actual resulting y — the value of the right-hand side of equation (B.4). This quantity is recorded on the y-axis. An equilibrium requires that the x and y coordinates are equal.

The line “ R and P short” plots y as if both R and P shorted, that is, their portfolio weights are calculated by adding the return φ to their perceived intrinsic expected return from the asset; in that case, the Sharpe ratio is given by (B.2). The line “Only P shorts” is produced similarly, except that the demand of agent R is set to zero; equation (B.3) applies. The actual resulting y is depicted by the thick continuous line, labeled “Actual response.” Finally, the line “Diagonal” depicts the equilibrium condition. Equilibria are

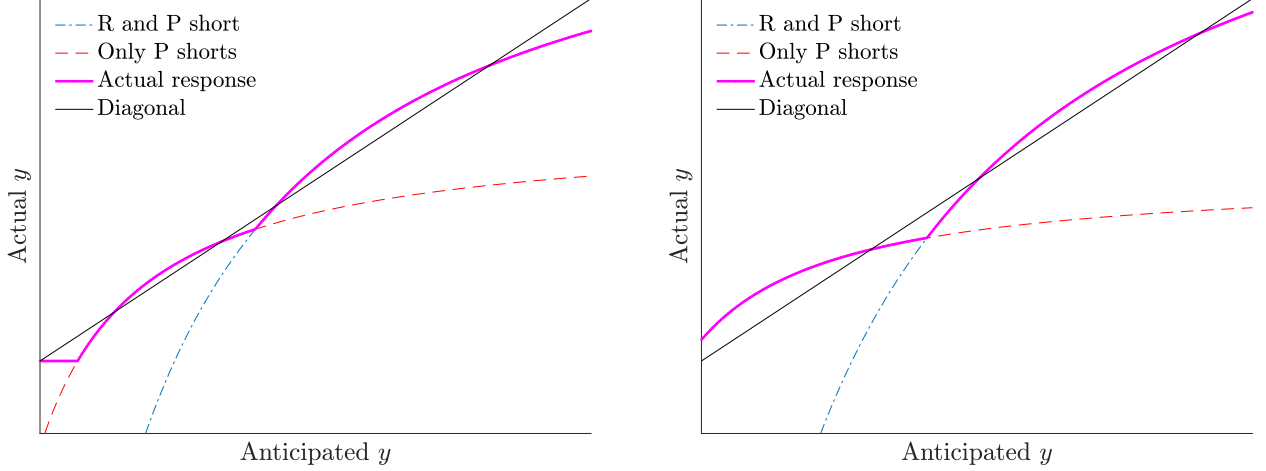


Figure B.1: The figure plots, in each panel, four lines pertaining to the model extension developed in this section. Equilibria are characterized by the satisfaction of equation (B.4), whose right-hand side is represented here by the line “Actual response” and the left-hand side by the line “Diagonal.” Further details are provided in the text.

therefore represented by points of intersection between the two continuous lines. The left panel presents a situation in which four equilibria with positive amounts of shorting and one with zero shorting obtain. The right panel presents a situation with three equilibria, all of which feature positive y .

We also flesh out the theoretical argument for the existence of a third equilibrium when ω^P is close to zero and two equilibria with $y > 0$ exist with $\omega^P = 0$ — i.e., the baseline model. By assumption, with $\omega^P = 0$ and $y = 0$ equation (B.3) applies and $\kappa + \frac{\varphi}{\sigma_D} > 0$. Choosing η^P so that $\kappa + \eta^P + \frac{\varphi}{\sigma_D} < 0$, we wish to conclude that equations

$$y = -\frac{\omega^P \left(\kappa + \eta^P + \frac{\varphi}{\sigma_D} \right)}{\omega^I \left(\kappa + \eta^I + \frac{\varphi}{\sigma_D} \tau y \right)} \quad (\text{B.5})$$

and (B.3) admit a solution that satisfies $\kappa + \frac{\varphi}{\sigma_D} > 0$ even for $\omega^P > 0$, at least when it is small enough. For simplicity, we keep ω^I constant as we increase ω^P from zero. Plugging (B.3) in (B.5) we obtain a quadratic that can be written as

$$y = \frac{\omega^P (\eta^I - \eta^P) \omega^I - \frac{\varphi}{\sigma_D} \omega^I \tau (1 - y) - \sigma_D}{\omega^I (\eta^I - \eta^P) \omega^P - \frac{\varphi}{\sigma_D} \omega^P \tau (1 - y) + \sigma_D} \equiv H(\omega^P, y). \quad (\text{B.6})$$

Our choice of η^P is such that the numerator of the second fraction on the right-hand side is positive at $y = 0$, which implies $\frac{\partial H}{\partial \omega^P} > 0$ evaluated at $\omega^P = 0$, as well as $\frac{\partial H}{\partial y} = 0$ at $\omega^P = 0$. We therefore have

$$\frac{dy}{d\omega^P} = \left(1 - \frac{\partial H}{\partial y} \right)^{-1} \frac{\partial H}{\partial \omega^P} > 0, \quad (\text{B.7})$$

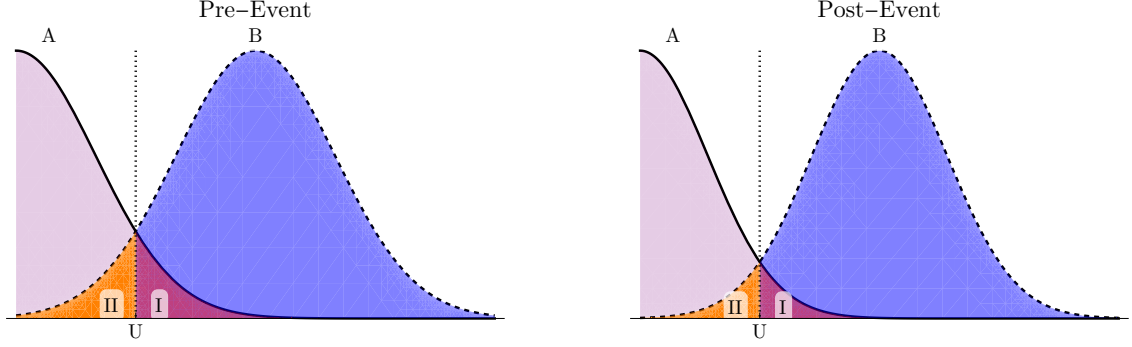


Figure C.1: Illustration of the effect of a decrease in volatility on the probability of Type-I and Type-II errors.

confirming that an equilibrium with positive y exists for small ω^P . (The condition $\kappa + \frac{\varphi}{\sigma_D} > 0$ is satisfied by continuity.)

C A Graphical Illustration of the Findings of Table 4.

In this section we provide a more detailed explanation for why a drop in volatility lowers the measured jump intensity of stocks that are unlikely to satisfy the condition of Proposition 4, but not for those that are likely to satisfy it.

For stocks that don't satisfy the condition of Proposition 4, the innovations to utilization are diffusive and the absolute value of the innovations to utilization (over discrete time-intervals) follows a distribution that resembles the half-normal distribution (since we are considering absolute values of utilization-changes.) This half-normal is depicted by distribution A in Figure C.1. Of course, since a jump is defined as a realization above U , there is the possibility of declaring a “false positive” jump with a probability that is depicted as Region I in the figure. Stocks that satisfy the condition of Proposition 4 could exhibit either positive or negative jumps in utilization. The relevant distribution of utilization-changes looks like the absolute value of a mixture distribution: If no jump occurs, the relevant distribution is A. If a jump occurs, then the relevant distribution is B. Distribution B is the same as A, but shifted to the right because of the jump. (Note that because we take absolute values, the shift is always to the right.) Since a “jump” is an observation above the cutoff U , there are possibilities of both false positive jumps (Region I), but also false negative jumps (Region II). These false negative jumps are instances where the jump has occurred, but the absolute value of the discrete-time change in utilization is still below the cutoff U .

The right plot of figure C.1 depicts a decline in the (diffusive) volatility of utilization. When the volatility falls, the area associated with Regions I and II shrinks.

The shrinkage of area I implies that the possibility of a false positive jump declines, thus lowering the (measured) jump rate of stocks that don't satisfy the condition of Proposition 4.

For stocks that satisfy the condition of Proposition 4 there are two opposing effects. On the one hand side, the shrinkage of area I lowers the instances of false positive jumps, which lowers the measured jump rate; on the other hand, the shrinkage of area II lowers the instances of false negatives, which raises the measured jump rate. In short, a decline in

volatility should unambiguously lower the measured jump rate for stocks that cannot exhibit jumps (fifth quintile stocks), but should have a lower (and in principle ambiguous effect) on the measured jump rate of stocks that are likely to satisfy the condition of Proposition 4 (first quintile stocks).

D Recursive Preferences with Unitary RRA

In this section we provide the details underlying the discussion in Section 6.1 in the text. For clarity, we develop the argument linearly, culminating in a formal result (Proposition 5).

We adopt the Epstein-Zin preference specification with the risk-aversion coefficient constrained to take a value of one:

$$V_t = E_t \int_t^\infty e^{-(\rho+\pi)(s-t)} \left(\frac{c^\alpha}{\alpha} ds - \frac{1}{2V_s} d[V_s] \right). \quad (\text{D.1})$$

This specification leads to V homogeneous of degree α in the consumption stream $\{c_t\}_t$, so that we can conjecture

$$V(W, x) = \frac{W^\alpha}{\alpha} g(x)^{\alpha-1}, \quad (\text{D.2})$$

where the function $g(\cdot)$, to be computed, will be shown below to equal the agent's consumption-to-wealth ratio.

Finally, we maintain the conjecture that an appropriate — and sufficient — state variable is the wealth share of agent R , denoted by ω . Because of unit risk aversion, the optimal portfolio coincides with the one obtained under log utility. Treating the volatility of returns, σ_t , as known, we make use of the known (myopic) portfolio policy to write

$$w^R = \frac{\kappa_t + \frac{f(y_t)}{\sigma_t}}{\sigma_t} 1_{\kappa_t + \frac{f(y_t)}{\sigma_t} < 0} + \frac{\kappa_t + \tau y_t \frac{f(y_t)}{\sigma_t}}{\sigma_t} 1_{\kappa_t + \tau y_t \frac{f(y_t)}{\sigma_t} > 0} \quad (\text{D.3})$$

$$w^I = \frac{\kappa_t + \eta + \frac{f(y_t)}{\sigma_t}}{\sigma_t} 1_{\kappa_t + \eta + \frac{f(y_t)}{\sigma_t} < 0} + \frac{\kappa_t + \eta + \tau y_t \frac{f(y_t)}{\sigma_t}}{\sigma_t} 1_{\kappa_t + \eta + \tau y_t \frac{f(y_t)}{\sigma_t} > 0}, \quad (\text{D.4})$$

noting further that the first term on the right-hand side of (D.4) equals zero in equilibrium, since I cannot short in equilibrium. We have the equilibrium restrictions

$$y_t = \frac{\omega_t}{1 - \omega_t} \frac{w^{R-}}{w^I} \quad (\text{D.5})$$

$$1 = w^R \omega_t + w^I (1 - \omega_t). \quad (\text{D.6})$$

A key determinant of the portfolio choices is the market return volatility σ_t , i.e., the

diffusion parameter in the equation

$$\begin{aligned} dR_t &= \frac{dS_t}{S_t} - \delta dt + \frac{D_t}{S_t} dt \\ &= \mu_t dt + \sigma_t dB_t. \end{aligned} \quad (\text{D.7})$$

We define the price-to-dividend ratio, $p_t = S_t/D_t$, and conjecture it to be a function of ω . The restriction that aggregate wealth equals the stock price and aggregate consumption the stock dividend gives

$$p(\omega) := \frac{S_t}{D_t} = \left(\sum g_i(\omega) \omega^i \right)^{-1}, \quad (\text{D.8})$$

so that we obtain

$$\sigma_t = \sigma_D + \frac{p'(\omega)}{p(\omega)} \sigma_\omega \quad (\text{D.9})$$

$$\mu_t = \mu_D - \delta + \frac{1}{p(\omega)} + \frac{p'(\omega)}{p(\omega)} \mu_\omega + \frac{1}{2} \frac{p''(\omega)}{p(\omega)} \sigma_\omega^2 + \frac{p'(\omega)}{p(\omega)} \sigma_\omega \sigma_D. \quad (\text{D.10})$$

We derive the dynamics of the wealth-share ω_t by remarking that $W_{R,t} = \omega_t S_t$ and thus

$$\frac{d(\omega S)}{\omega S} = (r + \pi + n) dt + w^R (dR_t + (f(y) 1_{w^R < 0} - r) dt) - g_R(\omega) dt - \pi dt + \frac{\nu \delta}{\omega} dt. \quad (\text{D.11})$$

Writing

$$d\omega_t = \mu_\omega dt + \sigma_\omega dB_t, \quad (\text{D.12})$$

we have the dynamics

$$\sigma_\omega + \omega \sigma_t = w^R \omega \sigma_t \quad (\text{D.13})$$

$$\mu_\omega + \omega(\mu_t + \delta - p(\omega)^{-1}) + \sigma_\omega \sigma_t = \omega (r_t + n + w^R (\mu_t + f(y) 1_{w^R < 0} - r_t) - g_R) + \nu \delta. \quad (\text{D.14})$$

Note that the drifts are subject to the agents' probability distortions. Specifically, agent I perceives $\mu_t^I = \mu_t + \eta \sigma_t$, which we already recognize in the portfolio choice problem. This must also be recognized in (D.14), i.e., agent I perceives

$$\mu_\omega^I = w^R \mu_t^I = \mu_\omega + \omega(w^R - 1) \eta \sigma_t = \mu_\omega + \eta \sigma_\omega. \quad (\text{D.15})$$

Equations (D.9) and (D.13) lead to

$$\sigma_t - \omega \frac{p'(\omega)}{p(\omega)} (w^R - 1) \sigma_t = \sigma_D. \quad (\text{D.16})$$

Given that the interest rate equals

$$r_t = \mu_t - \kappa_t \sigma_t, \quad (\text{D.17})$$

the system of equations (D.3)–(D.6) comprises four equations and has four unknowns — w^R , w^I , y , and κ — treating $g_i(\cdot)$ as known.

To solve this system together with (D.16), note that (D.16) gives σ_t^{-1} as linear in w^R :

$$\frac{1}{\sigma_t} = \frac{1}{\sigma_D} \left(1 + \frac{x p'(x)}{p(x)} (1 - w^R) \right). \quad (\text{D.18})$$

We can go through three cases, depending on the sign of w^R :

Case I: $w^R > 0$. We have

$$\kappa = \sigma_t - (1 - x)\eta \quad (\text{D.19})$$

$$w^R = \frac{\kappa}{\sigma_t} = 1 - \frac{(1 - x)\eta}{\sigma_t}. \quad (\text{D.20})$$

We can combine this equation with (D.18) to get

$$1 - w^R = (1 - x) \frac{\eta}{\sigma_D} (1 + x \log(p)' (1 - w^R)) \quad (\text{D.21})$$

$$1 - w^R = \left(1 - (1 - x)x \frac{\eta}{\sigma_D} \log(p)' \right)^{-1} (1 - x) \frac{\eta}{\sigma_D} \quad (\text{D.22})$$

$$w^R = 1 - \left(1 - (1 - x)x \frac{\eta}{\sigma_D} \log(p)' \right)^{-1} (1 - x) \frac{\eta}{\sigma_D}, \quad (\text{D.23})$$

which must be positive if it is to be equilibrium value for w^R .

If (D.23) gives w^R negative, then we may be in either Case II, in which the type R investor stays out of the market, or III, in which the type R investor shorts the stock.

Case II: $w^R = 0$. Here,

$$\kappa = \frac{\sigma_t}{1 - x} - \eta \quad (\text{D.24})$$

and we know from the violation of Case I that, with $w^R = 0$, the condition $\kappa < 0$ is satisfied. In order for this case to provide an equilibrium, it must also be the case that shorting, with $y = 0$, is unattractive:

$$\frac{1}{1 - x} - \frac{\eta}{\sigma_D} \left(1 + \frac{p'}{p} \right) + \frac{f(y)}{\sigma_D^2} \left(1 + \frac{p'}{p} \right)^2 \geq 0. \quad (\text{D.25})$$

The left-hand side of this inequality, too, uses the form σ_t takes when $w^R = 0$.

If this inequality is not satisfied, then we are (verified below) in

Case III: $w^R < 0$. We get, from the market-clearing condition,

$$\kappa = \sigma_t - \eta(1-x) - \frac{f(y)}{\sigma_t} (x + (1-x)\tau y), \quad (\text{D.26})$$

which we plug back in (D.3) to obtain

$$\begin{aligned} w^R &= 1 - \frac{\eta(1-x)}{\sigma_t} + \frac{f(y)}{\sigma_t^2} (1-x)(1-\tau y) \\ &= 1 - \frac{\eta(1-x)}{\sigma_D} \left(1 + (1-w^R) \frac{xp'}{p} \right) + \frac{f(y)}{\sigma_D^2} (1-x)(1-\tau y) \left(1 + (1-w^R) \frac{xp'}{p} \right)^2. \end{aligned} \quad (\text{D.27})$$

Finally, y is also expressed in terms of w^R :

$$y = -\frac{w^R x}{w^I(1-x)} = \frac{-w^R x}{1-w^R x}, \quad (\text{D.28})$$

leading to

$$\begin{aligned} w^R &= 1 - \frac{\eta(1-x)}{\sigma_D} \left(1 + (1-w^R) \frac{xp'}{p} \right) + \\ &\quad \frac{f(y)}{\sigma_D^2} (1-x) \left(1 + \tau \frac{w^R x}{1-w^R x} \right) \left(1 + (1-w^R) \frac{xp'}{p} \right)^2. \end{aligned} \quad (\text{D.29})$$

This is a cubic equation, whose solutions are available explicitly. For the sake of completeness, we write the cubic in full, as

$$0 = \sum_{i=0}^3 A_i^w w^i, \quad (\text{D.30})$$

with

$$A_3^w = -\frac{f(y)}{\sigma_D^2} (1-x) x^3 (1-\tau) \left(\frac{p'}{p} \right)^2 \quad (\text{D.31})$$

$$A_2^w = x - \frac{\eta}{\sigma_D} (1-x) x^2 \frac{p'}{p} + \frac{f(y)}{\sigma_D^2} (1-x) \left(2x(1-\tau) \left(1 + \frac{xp'}{p} \right) \frac{xp'}{p} + \left(\frac{xp'}{p} \right)^2 \right) \quad (\text{D.32})$$

$$A_1^w = -1 - x + \frac{\eta}{\sigma_D} (1-x) \left(\frac{xp'}{p} + x \left(1 + \frac{xp'}{p} \right) \right) - \quad (\text{D.33})$$

$$\frac{f(y)}{\sigma_D^2} (1-x) \left(x(1-\tau) \left(1 + \frac{xp'}{p} \right)^2 + 2 \left(1 + \frac{xp'}{p} \right) \frac{xp'}{p} \right)$$

$$A_0^w = 1 - \frac{\eta}{\sigma_D} (1-x) \left(1 + \frac{xp'}{p} \right) + \frac{f(y)}{\sigma_D^2} (1-x) \left(1 + \frac{xp'}{p} \right)^2. \quad (\text{D.34})$$

We also note explicitly the expression for κ that holds in all of the cases:

$$\kappa = \left(\frac{\sigma_t}{1-x} - \eta \right) (1_{w^R=0} + (1-x)1_{w^R>0}) + \left(\sigma_t - \eta(1-x) - \frac{f(y)}{\sigma_t} (x + (1-x)\tau y) \right) 1_{w^R<0}. \quad (\text{D.35})$$

We are ready to derive the ODE obeyed by the functions g_i . We write the Hamilton-Jacobi-Bellman (HJB) equation:

$$0 = \sup_{c,w} \frac{(cW)^\alpha}{\alpha} - \frac{1}{2\alpha} \frac{d[W^\alpha g^{\alpha-1}]_t}{W^\alpha g^{\alpha-1}} + \mathbb{E} \left[d \frac{e^{-(\rho+\pi)t} W^\alpha g^{\alpha-1}}{\alpha} \right]. \quad (\text{D.36})$$

This equation holds for each of the two agents, under their respective beliefs. Using the dynamics of W^i and $g(\omega_t)$, we expand (D.36) as

$$\begin{aligned} 0 = \sup_{c,w} \frac{(cW)^\alpha}{\alpha} - \frac{1}{2\alpha} \frac{d[W^\alpha g^{\alpha-1}]_t}{W^\alpha g^{\alpha-1}} + W^\alpha g^{\alpha-1} \left(r + n + \pi + w(\hat{\kappa}\sigma_t) - c - \frac{\rho + \pi}{\alpha} \right) + \\ \frac{\alpha-1}{\alpha} \frac{g'}{g} \mu_x W^\alpha g^{\alpha-1} + \frac{\alpha-1}{2} \left(w^2 \sigma_t^2 + 2w\sigma_t \frac{g'}{g} \sigma_x + \frac{(\alpha-2)}{\alpha} \left(\frac{g'}{g} \right)^2 \sigma_x^2 \right) W^\alpha g^{\alpha-1} + \\ \frac{\alpha-1}{2\alpha} \frac{g''}{g} \sigma_x^2 W^\alpha g^{\alpha-1}, \end{aligned} \quad (\text{D.37})$$

where the second term equals

$$\begin{aligned} \frac{1}{2\alpha} \frac{d[W^\alpha g^{\alpha-1}]_t}{W^\alpha g^{\alpha-1}} &= \frac{\alpha}{2} \left(w\sigma_t + \frac{\alpha-1}{\alpha} \frac{g'}{g} \sigma_x \right)^2 W^\alpha g^{\alpha-1} \\ &= \left(\frac{\alpha}{2} w^2 \sigma_t^2 + (\alpha-1)w\sigma_t \frac{g'}{g} \sigma_x + \frac{(\alpha-1)^2}{2\alpha} \left(\frac{g'}{g} \right)^2 \sigma_x^2 \right) W^\alpha g^{\alpha-1} \end{aligned} \quad (\text{D.38})$$

and we used the notation $\hat{\kappa}$ for $\hat{\kappa}^i$ defined as

$$\hat{\kappa}^i = \kappa + \eta^i + \frac{f(y)}{\sigma_t} 1_{w<0} + \frac{\tau y f(y)}{\sigma_t} 1_{w>0}. \quad (\text{D.39})$$

Plugging in the HJB, we have

$$\begin{aligned} 0 = \sup_{c,w} \frac{c^\alpha}{\alpha} + g^{\alpha-1} \left(r + n + \pi + w(\hat{\kappa}\sigma_t) - c - \frac{\rho + \pi}{\alpha} \right) - \frac{1}{2} w^2 \sigma_t^2 g^{\alpha-1} + \\ \frac{\alpha-1}{\alpha} \frac{g'}{g} \mu_x g^{\alpha-1} + \frac{\alpha-1}{2\alpha} \frac{g''}{g} \sigma_x^2 g^{\alpha-1} - \frac{\alpha-1}{2\alpha} \left(\frac{g'}{g} \right)^2 \sigma_x^2 g^{\alpha-1}, \end{aligned} \quad (\text{D.40})$$

which implies $c = g$ and $w = \sigma_t^{-1} \hat{\kappa}$ when $w \neq 0$ (interior solution).

Finally, we plug these values back in the HJB to obtain the ODE

$$0 = r + n + \pi + w^i \hat{\kappa}^i \sigma_t - \frac{1}{2} (w^i)^2 \sigma_t^2 - \frac{\rho + \pi}{\alpha} - \frac{\alpha - 1}{\alpha} g_i + \frac{\alpha - 1}{\alpha} \frac{g'_i}{g_i} \mu_\omega^i - \frac{\alpha - 1}{2\alpha} \left(\frac{g'_i}{g_i} \right)^2 \sigma_\omega^2 + \frac{\alpha - 1}{2\alpha} \frac{g''_i}{g_i} \sigma_\omega^2. \quad (\text{D.41})$$

In the above equation, the interest rate r is a function of the unknown functions g_i and their derivatives, as follows.

We start by computing

$$\frac{p'}{p} = - \left(\sum g_i(x) x_i \right)^{-1} \left(\sum g'_i(x) x_i + g_R - g_I \right) \quad (\text{D.42})$$

$$\begin{aligned} \frac{p''}{p} &= 2 \left(\sum g_i(x) x_i \right)^{-2} \left(\sum g'_i(x) x_i + g_R - g_I \right)^2 \\ &\quad - \left(\sum g_i(x) x_i \right)^{-1} \left(\sum g''_i(x) x_i + 2g'_R - 2g'_I \right) \end{aligned} \quad (\text{D.43})$$

and rewrite (D.10) as

$$\begin{aligned} \mu_t &= B_{x0} + B_{x1} \mu_x - \frac{1}{2} \left(\sum g_i(x) x_i \right)^{-1} \left(\sum g''_i(x) x_i \right) \sigma_x^2 \\ &= B_{x0} + B_{x1} \mu_x - B_{x2} \left(\sum g''_i(x) x_i \right) \end{aligned} \quad (\text{D.44})$$

for functions $B_{xk} = B_{xk}(g, g', x)$ and (D.14), based on the (myopic) optimal choice w^R , as

$$\begin{aligned} \mu_x + x(\mu_t + \delta - p(x)^{-1}) + \sigma_x \sigma_t &= x \left(\mu_t - \kappa \sigma_t + n + (\kappa + f(y) 1_{w^R < 0} / \sigma_t)^2 - g_R \right) + \nu \delta \\ \mu_x + x(\delta - p(x)^{-1}) + \sigma_x \sigma_t &= x \left(n - \kappa \sigma_t + (\kappa + f(y) 1_{w^R < 0} / \sigma_t)^2 - g_R \right) + \nu \delta. \end{aligned} \quad (\text{D.45})$$

Equation (D.45) gives μ_x explicitly. Then we have μ_t from (D.44), which we use in (D.17) to compute r explicitly. We have

$$r = \mu_D - \delta + \frac{1}{p(x)} + \frac{p'(x)}{p(x)} \mu_x + \frac{1}{2} \frac{p''(x)}{p(x)} \sigma_x^2 + \frac{p'(x)}{p(x)} \sigma_x \sigma_D - \kappa \sigma_t. \quad (\text{D.46})$$

A last element is the income flow due to the aggregate fees:

$$n = -(1 - \tau) f(y) w^R x 1_{w^R < 0}. \quad (\text{D.47})$$

We summarize our analysis in the following proposition.

Proposition 5 *A stationary equilibrium is characterized by functions g_i that obey the second-order system of coupled ODE (D.41) whose coefficients are defined as follows: p by (D.8), w^R by (D.23) or as a solution to (D.29), w^I by (D.6), y by (D.28), σ_t by (D.18), κ by (D.35), n by (D.47), the dynamics of ω by (D.14) and (D.13), with μ_ω^I from (D.15), and r by (D.46).*

An explicit computation of the equilibrium price requires a selection rule when multiple

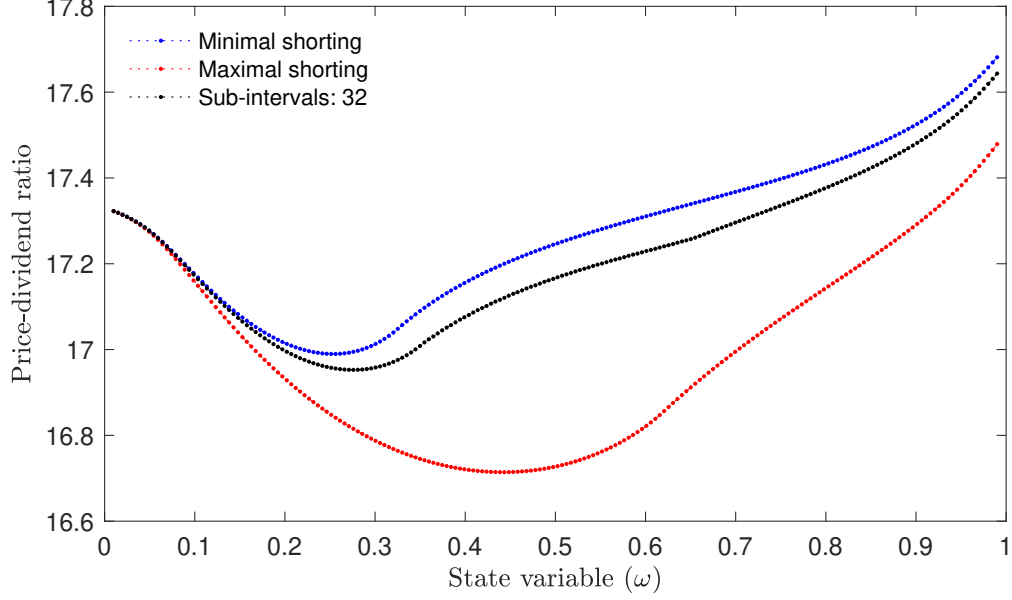


Figure D.1: Price-dividend ratios under different coordination assumptions in the multiplicity region $[\omega_1^*, \omega_2^*]$. The line “minimal shorting” assumes that investors always coordinate on the zero-shorting equilibrium whenever $\omega_t \in [\omega_1^*, \omega_2^*]$. The line “maximal shorting” assumes that investors coordinate on the highest shorting equilibrium whenever $\omega_t \in [\omega_1^*, \omega_2^*]$. The line “sub-intervals: 32” corresponds to splitting the interval $[\omega_1^*, \omega_2^*]$ into 32 sub-intervals and assumes high shorting on even-numbered sub-intervals and zero shorting on odd-numbered sub-intervals. The parameters are $\mu_D = 0.02$, $\sigma_D = 0.04$, $\pi = 0.05$, $\nu = 0.05$, $\delta = 0.05$, $\rho = 0$, $IES = 0.85$, and $\phi = 0.021$.

equilibria are possible. The lines “Maximal Shorting” and “Minimal Shorting” in Figure D.1 depict the price-dividend ratios obtaining when investors always coordinate on the maximal-shortening (resp. zero-shortening) equilibrium. The graph illustrates the general intuition we outlined in Section 6.1, explaining that an equilibrium selection featuring less shorting results in a higher price.

This model specification further enables us to address another question of interest, pertaining to the effect of the frequency of equilibrium switches. Specifically, one might fear that the expectation of switching back to a high-shortening equilibrium in the future reduces substantially the price response to the current switch to a low-shortening one.

To gain some insight in a tractable way, we implement a scheme of frequent equilibrium switching in the following way. We start by dividing the region of multiplicity (the interval $[\omega_1^*, \omega_2^*]$) into a relatively large number of equal-sized sub-intervals. On these sub-intervals, the coordination protocol prescribes high shorting on even-numbered sub-intervals and low shorting on odd-numbered sub-intervals. This specification allows us to capture the notion of frequent equilibrium switching, albeit in a stylized fashion.

The resulting price-dividend ratio is illustrated in Figure D.1 for the case of 32 sub-intervals. The figure shows that the lines labeled “Minimal shorting” and “Sub-intervals:32” are close, indicating that the price-dividend ratio with frequent equilibrium switches is close to the price-dividend ratio where investors always coordinate on the zero-shortening outcome. The reason is that, when the shorting market is active, the wealth share ω_t is more volatile.

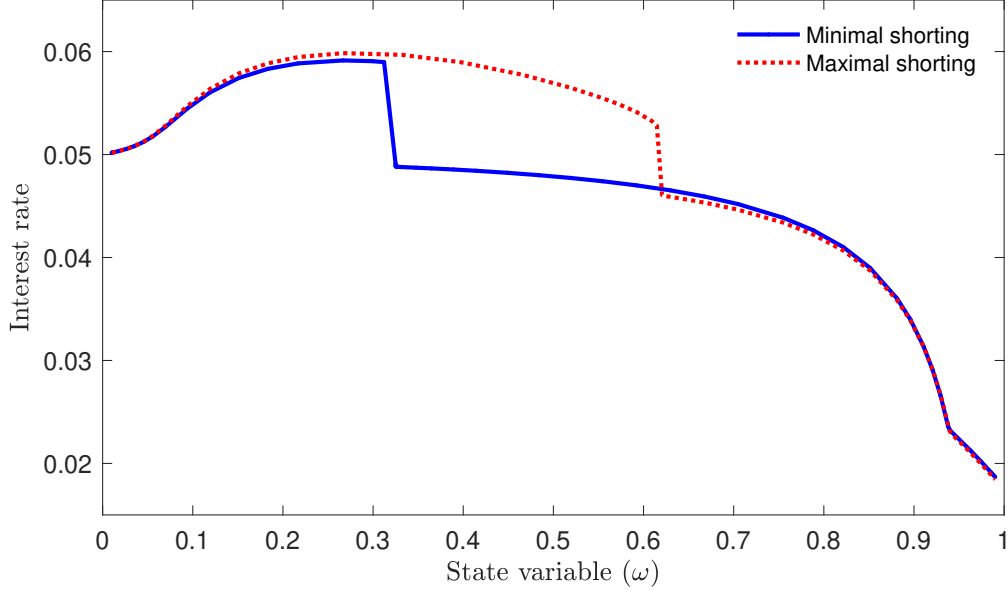


Figure D.2: Interest rate under different coordination assumptions in the multiplicity region $[\omega_1^*, \omega_2^*]$. The line “minimal shorting” assumes that investors always coordinate on the zero-shorting equilibrium whenever $\omega_t \in [\omega_1^*, \omega_2^*]$. The line “maximal shorting” assumes that investors coordinate on the highest shorting equilibrium whenever $\omega_t \in [\omega_1^*, \omega_2^*]$. The parameters are as for Figure D.1.

Accordingly, ω_t “travels fast” inside the regions where the shorting market is active and thus tends to exit these regions quickly. By contrast, when the shorting market is inactive, ω_t fluctuates less, moves more slowly, and as a result spends more time in these regions. We note that this is also the reason why the mass of the stationary distribution in Figure 3 of the paper is more concentrated in the region where the shorting market is inactive.

E A Two-Stock Economy with Limited Participation

In this section we introduce an additional security (stock 2) to our baseline model, which is not subject to any trading frictions. We continue to assume that borrowing stock 1 requires lending fees, as in the baseline model.

We allow one more feature, in the spirit of the “limited recognition hypothesis” of Merton (1987). Specifically, while all investors participate in the markets for stock 2 and the risk-free asset, only a fraction of investors pays any “attention” to stock 1. The remaining fraction of investors simply optimize their portfolio over the risk-free asset and stock 2 and assign zero weight to stock 1.

To ease the comparison of the results of this section with Proposition 2, we maintain the assumption that the lending supply curve is horizontal, and the lending fee is constant and equal to φ .

In this section, we take the equilibrium as given and derive several key restrictions that any equilibrium must obey; we solve fully for the equilibrium in Appendix F, in a limiting economy where stock 1 is small compared to stock 2. We assume that in equilibrium the

returns on stocks 1 and 2 follow a vector diffusion process of the form

$$dR_{1,t} = \mu_{1,t}dt + \sigma_{1,t}dB_{1,t} + b_t\sigma_{2,t}dB_{2,t} \quad (\text{E.1})$$

$$dR_{2,t} = \mu_{2,t}dt + \sigma_{2,t}dB_{2,t}, \quad (\text{E.2})$$

where $B_{1,t}$ and $B_{2,t}$ are independent Brownian motions, and $\mu_{1,t}$, $\mu_{2,t}$, $\sigma_{1,t}$, $\sigma_{2,t}$, and b_t are determined in equilibrium. We assume that investors I believe that Brownian motion 1 follows the dynamics⁵⁶ $dB_{1,t} + \eta dt$, while no investor has any belief distortions pertaining to Brownian motion 2.

To facilitate the statement of equilibrium returns, we define $\tilde{m}_{1,t} \equiv \frac{m_{1,t}}{\hat{\omega}_t}$ as the ratio of the stock-1 market capitalization share, denoted by $m_{1,t}$, to the wealth share of all investors participating in the market for stock 1, denoted by $\hat{\omega}_t$. We also define $\kappa_{1,t} \equiv \frac{(\mu_{1,t}-r)-b_t(\mu_{2,t}-r)}{\sigma_{1,t}}$ as the Sharpe ratio of a portfolio long 1 unit of asset 1 and short b_t units of asset 2.

Proposition 6 *In an equilibrium with shorting in asset 1 ($y_t > 0$), y_t is given by the root(s) of the quadratic equation*

$$0 = y \left(\eta + \frac{\tilde{m}_{1,t}}{\omega_t} \sigma_{1,t} - \frac{\varphi}{\sigma_{1,t}} (1 - \tau y) \right) - \left(\eta - \frac{\tilde{m}_{1,t}}{1 - \omega_t} \sigma_{1,t} - \frac{\varphi}{\sigma_{1,t}} (1 - \tau y) \right) \quad (\text{E.3})$$

that lie(s) in the interval $(0, 1)$, and the Sharpe ratio is given by

$$\kappa_{1,t} = \tilde{m}_{1,t} \sigma_{1,t} - (1 - \omega_t) \eta - \frac{\varphi}{\sigma_{1,t}} (\omega_t + (1 - \omega_t) \tau y_t). \quad (\text{E.4})$$

Similarly, in an equilibrium without shorting in asset 1 we have $\kappa_{1,t} = \sigma_{1,t} \tilde{m}_{1,t} - (1 - \omega_t) \eta$ if investor R holds an interior position in asset 1 and $\kappa_{1,t} = \frac{\sigma_{1,t} \tilde{m}_{1,t}}{1 - \omega_t} - \eta$ otherwise.

The excess return to asset 1 is given by

$$\mu_{1,t} - r_t = \kappa_{1,t} \sigma_{1,t} + b_t (\mu_{2,t} - r_t), \quad (\text{E.5})$$

where $\mu_{2,t} - r_t$ is the excess return of asset 2, given by the conventional CAPM relation

$$\mu_{2,t} - r_t = b_t \sigma_{2,t}^2 m_{1,t} + \sigma_{2,t}^2 m_{2,t}. \quad (\text{E.6})$$

Equations (E.3) and (E.4) are the same as (26) and (27), respectively, except that the volatility, σ_D , is replaced by $\tilde{m}_{1,t} \sigma_{1,t}$. The reason for this replacement is intuitive: In the case of a single stock, the risk of that stock, σ_D , is aggregate (by construction) and commands a risk premium. When there are multiple stocks, the risk compensation for bearing the idiosyncratic risk⁵⁷ of stock 1, $\sigma_{1,t}$, is multiplied by $\tilde{m}_{1,t}$, i.e., the stock market capitalization of stock 1 as a fraction of the wealth share of investors actively participating in the stock. An implication is that when $\tilde{m}_{1,t}$ approaches zero, the idiosyncratic risk becomes diversifiable,

56. More formally, the Radon-Nikodym derivative of the true probability measure with respect to the subjective one is given by $Z_t^I \equiv e^{-\frac{\eta^2}{2}t + \eta B_{1,t}}$.

57. Recall that the Sharpe ratio in Proposition 6 pertains to a portfolio that invests one dollar in asset 1 and shorts b_t units of asset 2, hedging out the exposure of the portfolio to the second Brownian shock.

and there is no compensation for bearing that risk (the first term on the right-hand side of (E.4) disappears).

Remark 3 *Since the equations determining $\kappa_{1,t}$ and y_t are essentially the same as (27) and (26), Proposition 4 remains unchanged when there are multiple stocks, except that now condition (b) becomes $(\tilde{m}_{1,t}\sigma_{1,t})^2 < \frac{1}{4}(1-y)^2|h'(y)|$. This condition therefore does not require that the total volatility of stock 1, or even its idiosyncratic part $\sigma_{1,t}$, be small, but rather that the risk of stock 1 be diversifiable by the agents trading it (small $\tilde{m}_{1,t}$).*

F The Price-Dividend Ratio of a Small Stock

This section provides the details of the entry-and-exit process for the model of Section 6.2.

We start with a few definitions. We let $\vec{m}_t$ denote the vector of market-capitalization weights of the two stocks, and $m_{j,t}$, $j \in \{1, 2\}$, its entries. Since the analysis of interest pertains to asset 1, from now on we use W_t^i to denote the wealth of all agents of type i that participate in the market for stock 1; the relevant state variable is $\omega_t^i \equiv W_t^i/(W_t^I + W_t^R)$, and to save notation we maintain the convention $\omega_t \equiv \omega_t^R$. As stated in the text, $\hat{\omega}_t$ denotes the wealth share of the investors who actively participate in the market for stock 1. We also let $\hat{w}_{2,t} = \frac{\mu_{2,t} - r_t}{\sigma_{2,t}^2}$ denote the optimal portfolio holding of stock 2 by investors who don't participate in stock 1, and $\vec{w}_t^i$ is the (row) vector of portfolio holdings of an investor $i \in \{I, R\}$ that is active in the market for stock 1. Finally, $\vec{B}_t \equiv (B_{1,t}, B_{2,t})^\top$.

We further define

$$\vec{B}_t \equiv \begin{bmatrix} B_{1,t} \\ B_{2,t} \end{bmatrix}, \quad \sigma_t = \begin{bmatrix} \sigma_{1,t} & b_t \sigma_{2,t} \\ 0 & \sigma_{2,t} \end{bmatrix}, \quad \vec{\varphi} = \begin{bmatrix} \varphi \\ 0 \end{bmatrix}, \quad \vec{\eta} = \begin{bmatrix} \eta \\ 0 \end{bmatrix}. \quad (\text{F.1})$$

The entry and exit into market 1 happens either for endogenous or exogenous reasons. By “endogenous” we mean that investors conduct a cost-benefit analysis before deciding whether to keep paying attention to the market for stock 1. In addition to this optimizing choice, we assume that investors enter and exit the market for exogenous reasons. This exogenous flux of investors is modeled with the sole purpose of making the model solution more tractable and transparent.

Specifically, with W_t^i the (aggregate) wealth of type- i investors that participate in market 1, we assume

$$dW_t^i = dW_t^{i,\text{part}} + \theta (\nu^i(W_t^I + W_t^R) - W_t^i) dt - 1_{i=R} \times \frac{W_t^I + W_t^R}{1 - \omega_t} dF_t + \omega_t^i (dL_t - dN_t), \quad (\text{F.2})$$

where $dW_t^{i,\text{part}}$ is the wealth growth of all investors of type $i \in \{I, R\}$ who already participate in the market for stock 1.⁵⁸ The term $\theta (\nu^i(W_t^I + W_t^R) - W_t^i) dt$ reflects entirely exogenous,

58. For completeness, $dW_t^{i,\text{part}} = W_t^{i,\text{part}} \mu_W^i dt + W_t^{i,\text{part}} (\vec{w}_t^i)^\top \sigma_t d\vec{B}_t$ where

$$\mu_W^i = r_t + \pi + n_t + (\vec{w}_{t,s}^i)^\top \left(\vec{\mu}_t - r_t \mathbf{1}_{2 \times 1} + \lambda_{t,s}^i \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) - \frac{c_{t,s}^i}{W_{t,s}^i}.$$

non-optimizing entry, which happens at some rate θ .

As in the baseline model, we are assuming that this exogenous entry-and-exit process affects the composition, but not the sum, of $W_t^I + W_t^R$, since

$$\sum_{i \in \{I, R\}} \theta (\nu^i (W_t^I + W_t^R) - W_t^i) = 0.$$

The term $-1_{i=R} \times \frac{W_t^I + W_t^R}{1 - \omega_t} dF_t$ captures the endogenous exit of R investors. As we described in the text, the (singular) process dF_t is constructed so that ω_t stays below the critical value $\bar{\omega}$ of ω_t (see (F.3) below) that ensures $V^R(\omega_t) > 0$ for $\omega_t < \bar{\omega}$.

Mostly for technical tractability reasons, we assume another source of exogenous entry and exit, which is reflected in the term $\omega_t^i (dL_t - dN_t)$ on the right-hand side of (F.2). This entry and exit process leaves the composition of wealth in the market (between R and I investors) unaffected, but ensures that the wealth of the investors who pay attention to the market 1 stays proportional to the “size” of market 1. Specifically, we define dL_t and dN_t as the two singular, increasing processes that control $W_t^I + W_t^R$ so that the ratio of stock market capitalization of asset 1 to the total wealth of investors participating in market 1, $\tilde{m}_t = \frac{M_{1,t}}{W_t^I + W_t^R}$, stays constant across time ($\tilde{m}_t = \tilde{m}$).⁵⁹ Because $(dL_t - dN_t)$ is multiplied by ω_t^i , this exogenous entry-and-exit process does not impact the composition of wealth between R and I investors. The purpose of this exogenous entry-and-exit term is transparency and tractability: By ensuring a constant $\tilde{m}_t$, if there were no differences of opinion ($\eta = 0$), the excess return, the price-dividend ratio, and the volatility of stock 1 would all be constant. Thus, we can eliminate a state variable from the problem, namely the ratio of market capitalization to the total wealth of investors in market 1. Economically, this means that we can abstract from the effects of limited participation (that have been studied extensively in the literature) and isolate the impact of shorting frictions. It is also worth highlighting that the term $\omega_t^i (dL_t - dN_t)$ endogenously approaches zero as δ_1 and θ approach infinity.⁶⁰ Thus, our computations would be approximately valid if we eliminated the term $\omega_t^i (dL_t - dN_t)$, as long as the analysis focuses on cases where investors are short-termist (θ is large) and the ratio of the dividends of a typical tree 1 to tree 2 mean reverts fast.

Having described the entry and exit of investors into the market for stock 1, we are ready to state a formal result describing the determination of equilibrium in this economy. For simplicity, we assume that the Brownian motions $B_{1,t}$ and $B_{2,t}$ are independent.

Proposition 7 *Using the expressions for w_t^i , $\kappa_{1,t}$ (with $b = 0$), and y_t from Proposition 6, the wealth share ω_t follows the diffusion process*

$$d\omega_t = \mu_\omega(\omega_t)dt + \sigma_\omega(\omega_t)dB_{1,t} - dF_t, \tag{F.3}$$

where F_t is an increasing (singular) process that reflects ω_t to remain below the value $\bar{\omega}$ that

⁵⁹. These processes can be uniquely constructed from the running maximum and minimum of the difference between $(W_t^R + W_t^I) - M_{1,t}$. For details see Karatzas and Shreve (2012, p. 210) on the Skorohod equation.

⁶⁰. The reason is that the price-dividend ratio and the ratio of the dividend processes for the two trees (given in (34)) approach constants, thus implying that $\tilde{m}_t$ approaches a constant ($\tilde{m}$).

is the lowest value for which $V^R(\omega_t) = 0$, and $\mu_\omega(\omega_t)$ and $\sigma_\omega(\omega_t)$ are given by

$$\mu_\omega(\omega_t) = \omega_t \left((w_{1,t}^R - \tilde{m}) \sigma_{1,t} (\kappa_t - \sigma_{1,t} \tilde{m}) + 1_{w_{1,t}^R < 0} w_{1,t}^R \varphi + \frac{y_t \tilde{m}}{1 - y_t} \varphi (1 - \tau) \right) + \quad (\text{F.4})$$

$$\theta (\nu - \omega_t),$$

$$\sigma_\omega(\omega_t) = \omega_t (w_{1,t}^R - \tilde{m}) \sigma_{1,t}, \quad (\text{F.5})$$

where $\sigma_{1,t} = \frac{p'(\omega_t)}{p(\omega_t)} \sigma_\omega(\omega_t) + \sigma_{1,D}$ is the volatility of stock 1 and the price-dividend ratio $p_t = p(\omega_t)$ solves the ordinary differential equation

$$\frac{1}{2} \frac{\partial^2 p}{\partial \omega_t^2} (\sigma_\omega(\omega_t))^2 + \frac{\partial p}{\partial \omega_t} (\mu_\omega(\omega_t) + (\sigma_{1,D} - \kappa_{1,t}) \sigma_\omega(\omega_t)) - p(r + \delta_1 + \kappa_{1,t} \sigma_{1,D}) + 1 = 0 \quad (\text{F.6})$$

in the region $0 \leq \omega_t \leq \bar{\omega}$.

Remark 4 Since there are multiple equilibrium values for w_t^i , $\kappa_{1,t}$, and y_t in Proposition 7, there exist a large set of solutions for $p(\cdot)$ and $\bar{\omega}$, depending on the equilibrium on which agents coordinate at each value of ω_t .

The expressions for μ_ω and σ_ω in Proposition 7 coincide with (30) and (29) when $\tilde{m} = 1$, $\theta = \pi$, and $\sigma_{1,t} = \sigma_D$.⁶¹ Moreover, with the dividend growths of stocks 1 and 2 being independent, so are their stock-price processes (in the limit where stock 1 becomes small) and the expressions for y_t , $w_{1,t}^i$, and $\kappa_{1,t}$ in Proposition 7 (with $\tilde{m} = 1$ and $\sigma_{1,t} = \sigma_D$) coincide with the respective expressions in the baseline model. Finally, if $\varepsilon = 0$, then $\bar{\omega} = 1$, as in the baseline model. In short, if one dropped the goods-market clearing requirement from the baseline model, the resulting expression for the price-to-dividend ratio would be given by (F.6) (with $\tilde{m} = 1$ and $\varepsilon = 0$).

The main complications with solving (F.6) are that a) it is a non-linear ODE⁶² and b) for $\varepsilon > 0$, this ODE is to be solved over a domain of values of ω_t on which $V^R(\omega_t) > 0$, with $V^R(\bar{\omega}) = 0$ and $\frac{dV^R(\bar{\omega})}{d\omega} = 0$ as boundary conditions.

We solve (F.6) with iterated use of MATLAB's ODE boundary value problem solver BVP5c. We start with the initial guess $\sigma_{1,t} = \sigma_{1,D}$. With that initial guess for $\sigma_{1,t}$ we solve the ODE for $V^R(\omega_t)$ using BVP5c for various boundaries $\bar{\omega}$ until we find the value of $\bar{\omega}$ that satisfies the boundary conditions $V^R(\bar{\omega}) = 0$ and $\frac{dV^R(\bar{\omega})}{d\omega} = 0$. With this $\bar{\omega}$ we compute the solution of the ordinary differential equation (F.6) on the interval $(0, \bar{\omega}]$ using the BVP5c solver and utilizing the (reflecting) boundary condition $p'(\bar{\omega}) = 0$. After obtaining the price-dividend ratio, $p(\omega_t)$ and its derivative, we evaluate $\frac{p'(\omega_t)}{p(\omega_t)}$, and compute $\sigma_{1,t} = \frac{p'(\omega_t)}{p(\omega_t)} \sigma_\omega(\omega_t) + \sigma_{1,D}$. Using this new guess for $\sigma_{1,t}$ we repeat the above procedure until convergence. For our numerical exercise, we assume that $\tilde{m} = 1$, to make the expressions for the Sharpe ratio, utilization, etc. directly comparable with the baseline model. The rest of the parameters are described in the text.

61. To see this, substitute the expression for the equilibrium interest rate (28) into (30).

62. Equation (F.6) is non-linear because μ_t^i and σ_t^i depend on $\sigma_{1,t}$, which in turn depends on $p(\cdot)$ and $p'(\cdot)$.

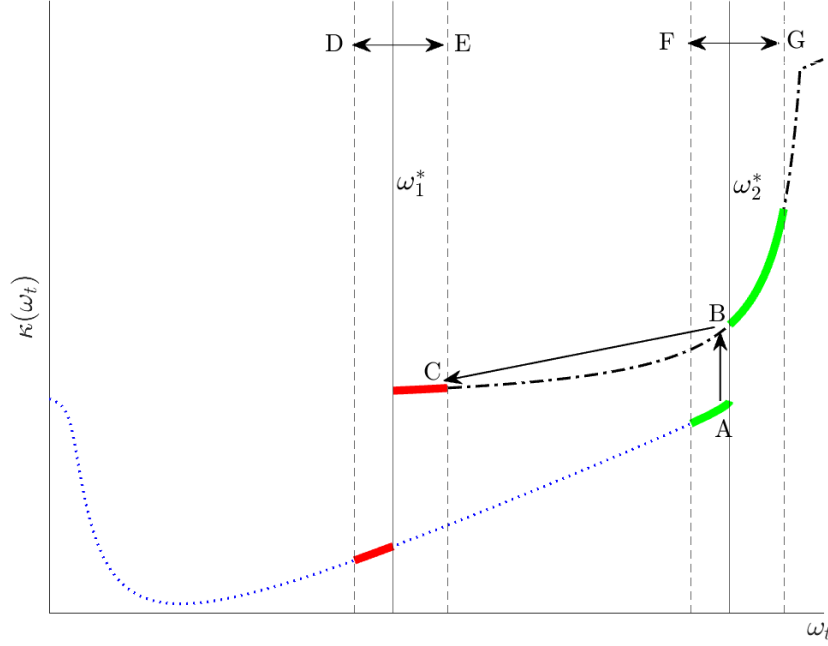


Figure F.1: An illustration of the short-run and long-run effects of an equilibrium shift.

F.1 The dependence of the price-dividend ratio on equilibrium coordination

To better explain how an equilibrium shift impacts the dynamics of the wealth shares and the Sharpe ratio, it is useful to refer to Figure F.1. The dash-dotted black line plots the Sharpe ratio values associated with the no shorting equilibrium. The blue dotted line shows the Sharpe ratios associated with the high shorting equilibrium. To the left of ω_1^* the Sharpe ratio is depicted by the blue line and is unique. Similarly, to the right of ω_2^* the Sharpe ratio is unique and depicted by the dash-dotted black line. Conditional on being in the high shorting equilibrium (point A), a shift from the high to the no shorting equilibrium can be decomposed into an “immediate” and an “eventual” effect. Figure F.1 illustrates the two effects. The first effect captures the immediate upward jump of the Sharpe ratio, and is depicted as a movement from point A to point B. We use the term “immediate” for this effect because the wealth shares are kept fixed. But the equilibrium shift also changes the dynamics of the wealth shares. As a result, the new typical stationary range of values of ω changes, and this is depicted as a move from point B to point C. In Figure F.1 the typical values of ω in the no shorting equilibrium are in the range D–E, while in the high shorting equilibrium, they are in the range F–G. When the market coordinates on the high shorting equilibrium, the associated typical values of the Sharpe ratio are the green points; similarly, when the market coordinates on the no-shorting equilibrium, the associated typical values of the Sharpe ratio are the red points. According to a Campbell-Shiller decomposition, the

log price-dividend ratio depends on the (geometrically-weighted) values of the risk premium over the infinite horizon. This average across the green points exceeds the respective average across the red points, and therefore the price-dividend ratio when agents coordinate on the high-shorting equilibrium is lower. (Intuitively, in the high shorting equilibrium the wealth shares places a significant mass to the right of ω_2^* , where the shorting market is inactive and the Sharpe ratio is particularly high.)

The small attention costs accelerate the transition from point A to point C. Upon an equilibrium shift to the no shorting equilibrium, a mass of short sellers endogenously chooses to exit the market for the small stock, thus making the transition from A to C essentially instantaneous.

G Adding a Non-Pecuniary Cost to Short Selling

In this section we make one change to the baseline model to allow for an additional — non-pecuniary — cost to shorting, meant to capture the increased burden of regulatory compliance brought about by the overhaul of Regulation SHO in 2008.

Specifically, suppose that the utility is given by

$$V_t^i = \int_t^\infty e^{-(\rho+\pi)(u-t)} \left(\log c_u^i + \frac{\chi}{\rho+\pi} w_u \times 1_{w_u < 0} \right) du, \quad (\text{G.1})$$

where $\frac{\chi}{\rho+\pi} w_u \times 1_{w_u < 0}$ captures this added regulatory burden.

As is intuitive, the higher cost to shorting depresses the demand for shorting, which in turn reduces portfolio heterogeneity across agents and therefore the variability of their relative wealths, as measured by $|\sigma_\omega|$. It follows that the volatility of the utilization ratio y also drops, too, at least for values of ω_t that are small enough.

We can state the following formal result.

Proposition 8 *(i) For ω_t such that the shorting market is active, the stable utilization ratio $y^+ > 0$ decreases with χ . (ii) The diffusive variance σ_ω^2 of ω_t decreases with χ . (iii) For values ω_t sufficiently low, the diffusive variance of y , denoted by σ_y^2 , decreases with χ .*

H Proofs

Proof of Proposition 1. Fix parameters $\eta > 0$ and $\psi > 1$ and define φ according to

$$\varphi = \sigma_D (\eta - \psi \sigma_D) \quad (\text{H.1})$$

for any value of σ_D . Note that when σ_D is sufficiently small, φ is guaranteed to be positive.

We show next that, as σ_D gets close to zero, Assumption 2 is satisfied. Rearranging (H.1) gives

$$\frac{\eta}{\frac{\varphi}{\sigma_D}} = \frac{1}{1 - \psi \frac{\sigma_D}{\eta}}. \quad (\text{H.2})$$

For sufficiently small σ_D we obtain

$$1 + \tau > \frac{1}{1 - \psi \frac{\sigma_D}{\eta}} > 1. \quad (\text{H.3})$$

Combining (H.2) and (H.3) yields (23).

Turning to (24), we note that the definition of ω_1^* along with (H.1) implies

$$\omega_1^* = 1 - \frac{\sigma_D}{\psi \sigma_D} = \frac{\psi - 1}{\psi} > 0,$$

while also

$$\lim_{\sigma_D \rightarrow 0} \frac{\sigma_D}{(1 + \tau) \frac{\varphi}{\sigma_D} - \eta} = \lim_{\sigma_D \rightarrow 0} \frac{\sigma_D}{(1 + \tau) (\eta - \psi \sigma_D) - \eta} = 0.$$

Therefore, for sufficiently small σ_D , the left-hand side of (24) converges to $\frac{\psi-1}{\psi} > 0$, while the right-hand side converges to zero, and therefore the inequality holds.

We conclude the proof by showing that $F(\omega)$ has a unique root in the interval $(\omega_1^*, 1)$. To this end, it is useful to introduce the definitions

$$A(\omega) \equiv \tau \frac{\omega}{\sigma_D} \varphi, \quad (\text{H.4})$$

$$B(\omega) \equiv \sigma_D - \omega \left((1 + \tau) \frac{\varphi}{\sigma_D} - \eta \right), \quad (\text{H.5})$$

$$C(\omega) \equiv \frac{\omega}{1 - \omega} \left(\sigma_D + (1 - \omega) \left(\frac{\varphi}{\sigma_D} - \eta \right) \right). \quad (\text{H.6})$$

With these definitions, $F(\omega)$ can be written as $F(\omega) = B^2(\omega) - 4A(\omega)C(\omega)$. We start by observing that $C(\omega_1^*) = 0$ for any parametric choice (since the definition of ω_1^* in equation (21) implies $\sigma_D + (1 - \omega_1^*) \left(\frac{\varphi}{\sigma_D} - \eta \right) = 0$). Also, Inequality (24) implies that $B(\omega_1^*) \neq 0$, and thus $B^2(\omega_1^*) > 0$. Accordingly, $F(\omega_1^*) > 0$. Also $B(1) < \infty$, while $C(1) = \infty$. By continuity, there exists at least one value $\omega_2^* \in (\omega_1^*, 1)$ such that $F(\omega_2^*) = 0$.

To show that this value is unique, consider any value $\omega_2^* \in (\omega_1^*, 1)$ such that $F(\omega_2^*) = 0$. We next show that $F'(\omega_2^*) < 0$.

To this end, note that

$$\begin{aligned} F'(\omega) &= 2B(\omega)B'(\omega) - 4[A'(\omega)C(\omega) + A(\omega)C'(\omega)] \\ &= 2B^2(\omega) \frac{B'(\omega)}{B(\omega)} - 4A(\omega)C(\omega) \left(\frac{A'(\omega)}{A(\omega)} + \frac{C'(\omega)}{C(\omega)} \right). \end{aligned}$$

Since ω_2^* is a root of $F(\omega)$ it follows that $B^2(\omega_2^*) = 4A(\omega_2^*)C(\omega_2^*)$. Therefore,

$$F'(\omega_2^*) = B^2(\omega_2^*) \left(2 \frac{B'(\omega_2^*)}{B(\omega_2^*)} - \frac{A'(\omega_2^*)}{A(\omega_2^*)} - \frac{C'(\omega_2^*)}{C(\omega_2^*)} \right). \quad (\text{H.7})$$

We have

$$\begin{aligned}\frac{A'(\omega_2^*)}{A(\omega_2^*)} &= \frac{1}{\omega_2^*} \\ \frac{B'(\omega_2^*)}{B(\omega_2^*)} &= -\frac{(1+\tau)\frac{\varphi}{\sigma_D} - \eta}{\sigma_D - \omega_2^* \left((1+\tau)\frac{\varphi}{\sigma_D} - \eta \right)}\end{aligned}$$

and

$$\frac{C'(\omega_2^*)}{C(\omega_2^*)} = \frac{1}{\omega_2^* (1 - \omega_2^*)} + \frac{\eta - \frac{\varphi}{\sigma_D}}{\sigma_D + (1 - \omega_2^*) \left(\frac{\varphi}{\sigma_D} - \eta \right)}.$$

Combining terms gives

$$\begin{aligned}2\frac{B'(\omega_2^*)}{B(\omega_2^*)} - \frac{A'(\omega_2^*)}{A(\omega_2^*)} - \frac{C'(\omega_2^*)}{C(\omega_2^*)} \\ = -\frac{2 \left((1+\tau)\frac{\varphi}{\sigma_D} - \eta \right)}{\sigma_D - \omega_2^* \left((1+\tau)\frac{\varphi}{\sigma_D} - \eta \right)} - \frac{1}{\omega_2^*} - \frac{1}{\omega_2^* (1 - \omega_2^*)} - \frac{\eta - \frac{\varphi}{\sigma_D}}{\sigma_D + (1 - \omega_2^*) \left(\frac{\varphi}{\sigma_D} - \eta \right)}.\end{aligned}\tag{H.8}$$

For future reference, we note that using $\omega_2^* > \omega_1^*$ along with (23) and the definition of ω_1^* implies that

$$\sigma_D + (1 - \omega_2^*) \left(\frac{\varphi}{\sigma_D} - \eta \right) > \sigma_D + (1 - \omega_1^*) \left(\frac{\varphi}{\sigma_D} - \eta \right) = 0.\tag{H.9}$$

Using (H.1) we can write the right-hand side of (H.8) as

$$-\frac{2 \left((1+\tau) (\eta - \psi \sigma_D) - \eta \right)}{\sigma_D - \omega_2^* \left((1+\tau) (\eta - \psi \sigma_D) - \eta \right)} - \frac{1}{\omega_2^*} - \frac{1}{\omega_2^* (1 - \omega_2^*)} - \frac{\psi}{1 - \psi (1 - \omega_2^*)}.\tag{H.10}$$

Taking the limit as σ_D approaches zero, the expression (H.10) converges to

$$-\frac{1}{1 - \omega_2^*} - \frac{\psi}{1 - \psi (1 - \omega_2^*)} < 0,$$

where the inequality follows from (H.9) along with (H.1).⁶³

The fact that the derivative $F'(\omega_2^*) < 0$ for any root of the equation $F(\omega_2^*) = 0$ in the interval $(\omega_1^*, 1)$ implies that the root ω_2^* must be unique. ■

Proof of Proposition 2. In preparation for the proof, we state and prove an auxiliary result.

63. Equation (H.1) implies $\frac{\varphi}{\sigma_D} - \eta = -\psi \sigma_D$, and therefore $0 < \sigma_D + (1 - \omega_2^*) \left(\frac{\varphi}{\sigma_D} - \eta \right) = \sigma_D (1 - (1 - \omega_2^*) \psi)$, where the inequality follows from (H.9).

Lemma 1 *The following statements hold for the quadratic equation (26).*

1. $\omega_1^* < \omega_2^*$ and the discriminant of (26) is non-negative for all $\omega_t \leq \omega_2^*$.
2. When $\omega_1^* \leq \omega_t \leq \omega_2^*$, the two roots of the equation are both in the interval $[0, 1]$.
3. For $\omega_t \in [0, \omega_1^*)$, only the larger root of (26) is in the interval $(0, 1)$.
4. If y is a root of (26), then $(1 - \omega_t)\eta - \sigma_D - \frac{1-\omega_t}{\sigma_D}\varphi(1 - \tau y) > 0$.

Proof of Lemma 1. We start with part 1. Using the definitions (H.4)–(H.6), equation (26) can be written in the familiar form

$$A(\omega_t)y^2 + B(\omega_t)y + C(\omega_t) = 0,$$

and the discriminant of this quadratic equation is given by $F(\omega_t)$ as defined in equation (22).

For $\omega_t \leq \omega_1^*$, $C(\omega_t) < 0$ and the discriminant, $B^2(\omega_t) - 4A(\omega_t)C(\omega_t)$, is positive. The assumption that ω_2^* is the unique root of $F(\omega)$ along with the facts that $F(\omega_1^*) = B^2(\omega_1^*) > 0$ and $F(1) = -\infty$ imply that $\omega_1^* < \omega_2^*$.⁶⁴ The uniqueness of the root ω_2^* also implies that $F(\omega_t) = B^2(\omega_t) - 4A(\omega_t)C(\omega_t) \geq 0$ for all $\omega_t \leq \omega_2^*$.

We now turn to part 2. To economize on notation we write A rather $A(\omega_t)$ and similarly for B and C . Fix a given ω_t and let $g(y) = Ay^2 + By + C$. We have $g(1) = A + B + C = \frac{\sigma_D}{1-\omega_t} > 0$ and $g'(1) = 2A + B = \sigma_D + \omega_t \left(\eta - (1 - \tau) \frac{\varphi}{\sigma_D} \right) > 0$, where the inequality follows from (23). Since $A > 0$, it follows that all roots of $g(y)$ must be smaller than one. Also, the fact that $\omega_t \geq \omega_1^*$ implies that $g(0) = C > 0$, while assumptions (23) and (24) together with the fact that $\omega_t \geq \omega_1^*$ imply that $g'(0) = B < 0$.

The facts that i) $g(y)$ is a convex, quadratic function of y , ii) $g(1) > 0$, $g(0) > 0$, $g'(1) > 0$, and $g'(0) < 0$ and iii) $B^2 - 4AC > 0$ for $\omega_t \in [\omega_1^*, \omega_2^*)$ imply that there are two roots in $(0, 1)$.

For part 3, we note that, when $\omega_t < \omega_1^*$, $g(0) = C < 0$, while $g(1) = A + B + C = \frac{\sigma_D}{1-\omega_t} > 0$. Therefore there exists one and only one root in $(0, 1)$.

Finally, let $y \in (0, 1)$ denote a root of the quadratic equation (26). Accordingly,

$$\begin{aligned} (1 - \omega_t)\eta - \sigma_D - (1 - \omega_t) \frac{\varphi}{\sigma_D} (1 - \tau y) &= \frac{1 - \omega_t}{\omega_t} y \left(\sigma_D + \omega_t \eta - \omega_t \frac{\varphi}{\sigma_D} (1 - \tau y) \right) \\ &= \frac{1 - \omega_t}{\omega_t} y \left(\sigma_D + \omega_t \left(\eta - \frac{\varphi}{\sigma_D} \right) + \omega_t \frac{\varphi}{\sigma_D} \tau y \right) \\ &> 0, \end{aligned}$$

where the last inequality follows from (23). This proves property 4. ■

We now continue with the proof of the proposition. We provide expressions for r_t and κ_t that apply in any equilibrium in which $w_t^R \neq 0$. Since $\sum_i \omega_t^i = 1$, it follows that $\sum_i \sigma_t^i = 0$

64. Assumption (24) implies that $B(\omega_1^*) \neq 0$ and therefore $B^2(\omega_1^*) > 0$.

and $\sum_i \mu_t^i = 0$. Using (29) and $\sum_i \sigma_t^i = 0$ implies that $\sum_i \omega_t^i w_t^i = 1$. Combining $\sum_i \omega_t^i w_t^i = 1$ with (12) along with the definition $y_t = \frac{W_t^-}{W_t^+}$ gives

$$\kappa_t + (1 - \omega_t) \eta + \left(\omega_t \frac{1}{\sigma_D} \varphi + (1 - \omega_t) \tau y_t \frac{1}{\sigma_D} \varphi \right) 1_{\{w_t^R < 0\}} = \sigma^D. \quad (\text{H.11})$$

Similarly, using (30) along with $\sum_i \mu_t^i = 0$ and $\sum_i \omega_t^i (n_t + w_t^i s_t^i) = 0$ gives (28).

We next describe the equilibria for the three intervals of ω_t described in the statement of the proposition.

i) In this case, $\omega_t > \omega_2^*$. The equilibrium prescribes non-negative portfolios for both investors. If $\omega_t > 1 - \frac{\sigma_D}{\eta}$, equation (H.11) implies that $\kappa_t > 0$ and (12) implies that both investors hold positive portfolios and the shorting market is inactive. If $\omega_t \in [\omega_1^*, 1 - \frac{\sigma_D}{\eta})$, then there exists an equilibrium that involves no shorting and a zero portfolio for investor R . We check this assertion by observing that the associated market clearing requirement becomes $(1 - \omega_t) w_t^I = 1$, which together with $y_t = 0$ leads to (25). We then note that

$$\begin{aligned} \kappa_t + \frac{\varphi}{\sigma_D} &= \frac{\sigma_D}{1 - \omega_t} - \eta + \frac{\varphi}{\sigma_D} \\ &> \frac{\sigma_D}{1 - \omega_1^*} - \eta + \frac{\varphi}{\sigma_D} \\ &= 0. \end{aligned} \quad (\text{H.12})$$

The first line follows from (25), the second line follows from $\omega_t > \omega_1^*$ and the third line follows from the definition of ω_1^* . Since $\kappa_t + \frac{\varphi}{\sigma_D} > 0$, investor R does not choose a negative portfolio. And since $\kappa_t < 0$ for $\omega_t \in [\omega_1^*, 1 - \frac{\sigma_D}{\eta})$, the investor chooses a zero portfolio.

ii) In this case, $\omega_1^* < \omega_t < \omega_2^*$. Since $\omega_t > \omega_1^*$, equation (H.12) implies that the no-shorting equilibrium continues to be an equilibrium. There exist, however, two more equilibria. To compute them, we guess (and verify shortly) that $w_t^R < 0$. Using (12) and (H.11) gives

$$\begin{aligned} y_t &= \frac{W_t^-}{W_t^+} = \frac{-\omega_t w_{t,s}^R}{(1 - \omega_t) w_{t,s}^I} = \frac{\omega_t}{1 - \omega_t} \frac{-\left(\kappa_t + \frac{1}{\sigma_D} \varphi\right)}{\kappa_t + \eta_t + \frac{1}{\sigma_D} \varphi \tau y_t} \\ &= \frac{\omega_t}{1 - \omega_t} \frac{(1 - \omega_t) \eta - \sigma_D - \frac{1 - \omega_t}{\sigma_D} \varphi (1 - \tau y_t)}{\sigma_D + \omega_t \eta - \frac{\omega_t}{\sigma_D} \varphi (1 - \tau y_t)}. \end{aligned}$$

Rearranging leads to (26). Statement 1 of Lemma 1 implies that, when $\omega_t \in (\omega_1^*, \omega_2^*)$, equation (26) has two roots in $(0, 1)$. Under the supposition that $w_t^R < 0$, equation (H.11) leads to (27). In turn

$$\begin{aligned} \kappa_t^\pm + \frac{\varphi}{\sigma_D} &= \sigma_D - (1 - \omega_t) \eta - \frac{\omega_t}{\sigma_D} \varphi \left(1 + \tau y^\pm \frac{1 - \omega_t}{\omega_t} \right) + \frac{\varphi}{\sigma_D} \\ &= \sigma_D - (1 - \omega_t) \left(\eta + \frac{\varphi}{\sigma_D} (1 - \tau y_t^\pm) \right) < 0, \end{aligned} \quad (\text{H.13})$$

where the last inequality follows from statement 4 of Lemma 1. Combining this observation

with (12) confirms that $w_t^R < 0$. Note that in the second and third equilibria we have that

$$\kappa_t^\pm + \eta_t + \frac{1}{\sigma_D} \varphi \tau y_t^\pm = \sigma_D + \omega_t \eta - \frac{\varphi \omega_t}{\sigma_D} (1 - \tau y_t^\pm) > 0,$$

where the last inequality follows from (H.13) along with the fact that $y^\pm$ satisfy the equation (26). This implies that $w_t^I > 0$.

iii) In this case, $\omega_t < \omega_1^*$. Statement 3 of Lemma 1 implies that the quadratic equation (26) has only one solution in $(0, 1)$. This shows that there can only be one equilibrium with shorting. Moreover, this is the unique equilibrium. If w_t^R were zero and the Sharpe ratio were $\frac{\sigma_D}{1-\omega_t} - \eta$, then the inequality in (H.12) reverses, i.e., $\frac{\sigma_D}{1-\omega_t} - \eta + \frac{\varphi}{\sigma_D} < 0$ and investor R would want to deviate from the equilibrium prescription and choose a negative portfolio.

The dynamics of the wealth share follow from a straightforward application of Ito's lemma.

■

Lemma 2 *When the equilibrium is unique, $0 < \Phi < 1$.*

Proof of Lemma 2. We start by noting that an application of the implicit function theorem to (26) gives $\frac{dy}{d\eta} = \frac{1-y}{Z'(y)}$, where $Z(y) \equiv y \left(\eta + \frac{\sigma_D}{\omega_t} - \frac{\varphi}{\sigma_D} (1 - \tau y) \right) - \left(\eta - \frac{\sigma_D}{1-\omega_t} - \frac{\varphi}{\sigma_D} (1 - \tau y) \right)$. $Z(y)$ is a quadratic equation in y with positive leading coefficient, and satisfies $Z(0) < 0$ when $\omega_t < \omega_1^*$. There consequently exists a unique value $y > 0$ such that $Z(y) = 0$; for this value, $Z'(y) > 0$. Hence, $\frac{dy}{d\eta} > 0$.

Next note that $G_y = \kappa + \eta + 2\frac{\varphi}{\sigma_D} \tau y > 0$, $G_\kappa = y + \frac{\omega_t}{1-\omega_t} > 0$, and $G_\eta = y > 0$. This proves $\Phi > 0$.

Finally, note that $Z(y) = G(y, \kappa(y))$. Therefore, $Z'(y) = G_y + G_\kappa \frac{d\kappa}{dy} = G_y (1 - \Phi)$. Since $Z'(y) > 0$ at the equilibrium value of y , it follows that $G_y (1 - \Phi) > 0$. Since $G_y = y > 0$, it follows that $\Phi < 1$. ■

Proof of Proposition 3. We note first that, for $w \leq 0$, the function

$$\iota(w, \kappa) \equiv w (\kappa \sigma_D + \varphi) - \frac{1}{2} (w \sigma_D)^2 \tag{H.14}$$

is decreasing in κ , and therefore it attains a higher maximum for equilibrium B (since $\kappa^B < \kappa^A$).

It immediately follows that

$$g_t^B - g_t^A = -(\kappa_t^B - \kappa_t^A) \sigma_D + \max_{w \leq 0} \iota(w, \kappa_t^B) - \max_{w \leq 0} \iota(w, \kappa_t^A) \geq 0.$$

We further have, based on the expressions for g_t and μ_ω (equation (30)),

$$\begin{aligned} \mu_\omega^B(\omega_t) - \mu_\omega^A(\omega_t) &= \omega_t (g_t^B - g_t^A) + \frac{1}{2} \omega_t (w_t^B (w_t^B - 2) - w_t^A (w_t^A - 2)) \sigma_D^2 \\ &= \omega_t (g_t^B - g_t^A) + \frac{1}{2} \omega_t (w_t^B - w_t^A) (w_t^B + w_t^A - 2) \sigma_D^2 \\ &> 0, \end{aligned}$$

with both of the factors in parentheses in the second term on the right-hand side of the second-to-last line being negative (since $w_t^B < w_t^A < 0$). ■

Proof of Proposition 4. We start by describing the determination of the equilibrium in this case. Fix a time t and let E denote expectations with respect to the wealth distribution over types η at time t . (For notational simplicity, we remove time-subscripts throughout the proof.) For a given Sharpe ratio κ and anticipated utilization ration y , define the following two functions, giving the aggregate long and short positions, respectively.

$$L(y, \kappa) = E \left[\sigma^{-1} (\eta + \kappa + \sigma^{-1} \tau y f(y))^+ \right] \quad (\text{H.15})$$

$$S(y, \kappa) = E \left[\sigma^{-1} (\eta + \kappa + \sigma^{-1} f(y))^- \right]. \quad (\text{H.16})$$

An equilibrium is defined through the two market-clearing conditions

$$1 = L(y, \kappa) - S(y, \kappa) \quad (\text{H.17})$$

$$y = \frac{S(y, \kappa)}{L(y, \kappa)}. \quad (\text{H.18})$$

Furthermore, (H.17) defines κ uniquely as a function of y , so that we can write $S(y) = S(y, \kappa(y))$ and $L(y) = L(y, \kappa(y))$, and the equilibrium determination comes down to

$$F(y) \equiv \frac{S(y)}{L(y)} = y. \quad (\text{H.19})$$

The remainder of the proof is organized as follows. We start by showing that, given y_1 with $h'(y_1) < 0$, a continuous distribution with connected support (thus the density does not drop to zero on an intermediate range to then become positive again) exists for which $F'(y_1) > 1$. Using this property, we show that there exist multiple equilibria for this distribution. The continuity of the problem then ensures that, for any sequence of distributions converging to the one we construct,⁶⁵ a sequence of equilibrium utilization rates $y_1^{(n)}$ obtain that converges to y_1 , and consequently $F'(y_1^{(n)}) > 1$ for n large enough. In this sense, the set of type distributions admitting multiple equilibria is not “knife-edge” or even sparse, but in fact has non-empty interior.

For convenience, we define $\bar{h}(y) = \frac{h(y)}{\sigma}$ and note that $\bar{h}'(y) < 0$ is equivalent to $h'(y) < 0$. Equation (H.17) implies that

$$\kappa(y) = \frac{\sigma - \omega^S \bar{\eta}^S - \omega^L \bar{\eta}^L - \left(\omega^S \frac{f(y)}{\sigma} + \omega^L \tau y \frac{f(y)}{\sigma} \right)}{\omega^S + \omega^L}, \quad (\text{H.20})$$

65. Convergence in the space of distribution is defined in terms of convergence of expectations of any smooth function with compact support.

where we defined the quantities

$$\omega^L = \mathbb{E} [1_{\{\eta + \kappa + \sigma^{-1} \tau y f(y) \geq 0\}}] \quad (\text{H.21})$$

$$\omega^S = \mathbb{E} [1_{\{\eta + \kappa + \sigma^{-1} f(y) \leq 0\}}] \quad (\text{H.22})$$

$$\bar{\eta}^L = \mathbb{E} [\eta \mid \eta + \kappa + \sigma^{-1} \tau y f(y) \geq 0] \quad (\text{H.23})$$

$$\bar{\eta}^S = \mathbb{E} [\eta \mid \eta + \kappa + \sigma^{-1} f(y) \leq 0]. \quad (\text{H.24})$$

(These quantities depend on y , but we suppress that dependence in our notation.)

Furthermore, one can differentiate the same equation (H.17) with respect to y to obtain

$$\kappa'(y) = -\sigma^{-1} \frac{\omega^S f'(y) + \tau \omega^L (f(y) + y f'(y))}{\omega^S + \omega^L}, \quad (\text{H.25})$$

where we have made use of the fact that $\frac{d}{dx} \mathbb{E}[(g(x, \eta))^+] = \mathbb{E} \left[\frac{d}{dx} g(x, \eta) 1_{\{g(x, \eta) \geq 0\}} \right]$ for an arbitrary differentiable function g , given that the distribution of η is absolutely continuous.

Using equations (H.16) and (H.20) and the definitions of $h(y)$ and $\bar{h}(y)$, we compute

$$S(y) = \sigma^{-1} \frac{\omega^L \omega^S}{\omega^L + \omega^S} \left(\bar{\eta}^L - \bar{\eta}^S - \frac{\sigma}{\omega^L} - \bar{h}(y) \right) = B^{-1} (A - \bar{h}(y)) \quad (\text{H.26})$$

$$F(y) = \frac{\bar{\eta}^L - \bar{\eta}^S - \frac{\sigma}{\omega^L} - \bar{h}(y)}{\bar{\eta}^L - \bar{\eta}^S + \frac{\sigma}{\omega^S} - \bar{h}(y)} = \frac{A - \bar{h}(y)}{A + B - \bar{h}(y)}, \quad (\text{H.27})$$

where we also defined

$$A \equiv \bar{\eta}^L - \bar{\eta}^S - \frac{\sigma}{\omega^L} \quad (\text{H.28})$$

$$B \equiv \frac{\sigma}{\omega^S} + \frac{\sigma}{\omega^L}. \quad (\text{H.29})$$

Noting now, using (H.16) and (H.25), that

$$S'(y) = -\sigma^{-1} \frac{\omega^L \omega^S}{\omega^S + \omega^L} \bar{h}'(y) = -B^{-1} \bar{h}'(y), \quad (\text{H.30})$$

we use (H.26) and (H.30), as well as $F(y) = \frac{S(y)}{L(y)} = \frac{S(y)}{1+S(y)} = 1 - \frac{1}{1+S(y)}$, to write

$$F'(y) = \frac{S'(y)}{(1+S(y))^2} = \frac{-B \bar{h}'(y)}{(A+B-\bar{h}'(y))^2}. \quad (\text{H.31})$$

Our intermediate goal, therefore, is to show that, given $\bar{h}'(y_1) < 0$, values A and B exist

satisfying

$$\frac{A - \bar{h}(y_1)}{A + B - \bar{h}(y_1)} = y_1 \quad (\text{H.32})$$

$$\frac{-B\bar{h}'(y_1)}{(A + B - \bar{h}(y_1))^2} = 1 + \varepsilon > 1 \quad (\text{H.33})$$

for some $\varepsilon > 0$. In fact, for any $\varepsilon > 0$, solutions A and B to these equations are given by

$$B = (1 - y_1)^2 \frac{|\bar{h}'(y_1)|}{1 + \varepsilon} > 0 \quad (\text{H.34})$$

$$A = \bar{h}(y_1) + B \frac{y_1}{1 - y_1} = \bar{h}(y_1) + y_1(1 - y_1) \frac{|\bar{h}'(y_1)|}{1 + \varepsilon} > \bar{h}(y_1). \quad (\text{H.35})$$

To show the existence of a distribution yielding these desired values of A and B , we first note that the right-hand side of (H.29) can be made arbitrarily close to 4σ while keeping $\omega^L + \omega^S < 1$, and therefore condition b) of the proposition ensures that such ω^L and ω^S exist delivering B for a small enough ε . Fixing ω^L and ω^S , $\bar{\eta}^L$ and $\bar{\eta}^S$ can be chosen arbitrarily subject to (H.28) delivering the desired value of A . We therefore now have the value of $\kappa(y_1)$, which determines the sets of types that go long, respectively short, the asset. Finally, the density of the distribution on each of these two sets can be chosen freely subject to the two integrals defining ω^L and $\bar{\eta}^L$, respectively ω^S and $\bar{\eta}^S$. In the complementary, intermediate type region in which agents are inactive, the density is only subject to a total mass condition.

Finally, with $Y = \min\{1, y | \bar{h}(y) = A\}$, either $Y < 1$ and $F(Y) = 0 < Y$ or $F(Y) = F(1) < 1 = Y$. Since $F(Y) < Y$ in either case, and $F'(y_1) > 1$, a value $y_2 \in (y_1, Y)$ exists such that $y_2 = F(y_2)$. Thus, a second equilibrium exists. ■

Proof of Proposition 6. The proof essentially repeats the steps from the one-risky asset case, so we provide only a sketch, focusing on the elements that differ.

With these definitions, the market clearing condition is

$$\hat{\omega}_t \sum_{i \in \{I, R\}} \omega_t^i \vec{w}_t^i + (1 - \hat{\omega}_t) \begin{bmatrix} 0 \\ \hat{w}_{2,t} \end{bmatrix} = \vec{m}_t. \quad (\text{H.36})$$

We consider first an equilibrium with $y_t > 0$. Investor R 's and I 's optimal portfolios are given by

$$\vec{w}_t^R = (\sigma_t \sigma_t')^{-1} (\vec{\mu}_t - r_t \mathbf{1}_{2 \times 1} + \vec{\varphi}), \quad (\text{H.37})$$

$$\vec{w}_t^I = (\sigma_t \sigma_t')^{-1} (\vec{\mu}_t - r_t \mathbf{1}_{2 \times 1} + \sigma_{1,t} \vec{\eta} + \tau y_t \vec{\varphi}). \quad (\text{H.38})$$

Using (H.37) inside (H.36) yields

$$\begin{aligned}
(\sigma_t \sigma'_t) \vec{m}_t &= \widehat{\omega}_t (\omega_t (\vec{\mu}_t - r \mathbf{1}_N + \vec{\varphi}) + (1 - \omega_t) (\vec{\mu}_t - r \mathbf{1}_N + \sigma_1 \vec{\eta} + \tau y_t \vec{\varphi})) \\
&\quad + (1 - \widehat{\omega}_t) (\sigma_t \sigma'_t) \begin{bmatrix} 0 \\ \frac{\mu_{2,t} - r}{\sigma_{2,t}^2} \end{bmatrix}.
\end{aligned} \tag{H.39}$$

Next we use the row selection vector $[0, 1]$ to pre-multiply both sides of (H.39). Noting that $[0, 1] \vec{\varphi} = [0, 1] \vec{\eta} = 0$, and also

$$(\sigma_t \sigma'_t) \begin{bmatrix} 0 \\ \frac{\mu_{2,t} - r}{\sigma_{2,t}^2} \end{bmatrix} = \begin{bmatrix} b_t (\mu_{2,t} - r) \\ \mu_{2,t} - r \end{bmatrix}, \tag{H.40}$$

leads to (E.6). We next note that

$$\begin{aligned}
[1, -b_t] \sigma_t \sigma'_t \begin{bmatrix} m_{1,t} \\ m_{2,t} \end{bmatrix} &= [\sigma_{1,t}, 0] \begin{bmatrix} \sigma_{1,t} & 0 \\ b_t \sigma_{2,t} & \sigma_{2,t} \end{bmatrix} \begin{bmatrix} m_{1,t} \\ m_{2,t} \end{bmatrix} \\
&= \sigma_{1,t}^2 m_{1,t}.
\end{aligned} \tag{H.41}$$

Pre-multiplying both sides of (H.39) with the row vector $[1, -b_t]$, using (H.40), (H.41), and the definition of $\kappa_{1,t}$, and re-arranging yields

$$\kappa_{1,t} = \widetilde{m}_{1,t} \sigma_{1,t} - (1 - \omega_t) \eta - \frac{\varphi}{\sigma_{1,t}} (\omega_t + (1 - \omega_t) \tau y_t). \tag{H.42}$$

Using the definition of $\kappa_{1,t}$ inside (H.37) gives

$$w_{1,t}^R = \frac{\kappa_{1,t}}{\sigma_{1,t}} + \frac{\varphi}{\sigma_{1,t}^2} \tag{H.43}$$

$$w_{1,t}^I = \frac{\kappa_{1,t} + \eta}{\sigma_{1,t}} + \frac{\tau y_t \varphi}{\sigma_{1,t}^2}, \tag{H.44}$$

where we used the notation $w_{1,t}^i$, $i \in \{R, I\}$, to denote the first element of w_t^i .

Using the market clearing condition $y_t = -\frac{\omega_t^R w_{1,t}^R}{\omega_t^I w_{1,t}^I} = -\frac{\omega_t w_{1,t}^R}{(1 - \omega_t) w_{1,t}^I}$ leads to (E.3).

If agent R chooses not to short then the market clearing condition becomes

$$\widehat{\omega}_t (1 - \omega_t) \vec{w}_t^I + (1 - \widehat{\omega}_t) \begin{bmatrix} 0 \\ \widehat{w}_{2,t} \end{bmatrix} = \vec{m}_t. \tag{H.45}$$

Substituting in $\vec{w}_t^I$ from (H.38) and pre-multiplying by $\sigma_t \sigma'_t$ gives

$$(\sigma_t \sigma'_t) \vec{m}_t = \widehat{\omega}_t (1 - \omega_t) (\vec{\mu}_t - r \mathbf{1}_N + \sigma_1 \vec{\eta}) + (1 - \widehat{\omega}_t) (\sigma_t \sigma'_t) \begin{bmatrix} 0 \\ \frac{\mu_{2,t} - r}{\sigma_{2,t}^2} \end{bmatrix}. \tag{H.46}$$

Premultiplying (H.46) by the row $[1, -b_t]$ and using (H.40) and (H.41) gives

$$\sigma_{1,t}^2 \tilde{m}_{1,t} = (1 - \omega_t) \sigma_{1,t} (\kappa_{1,t} + \eta),$$

and therefore

$$\kappa_{1,t} = \sigma_{1,t} \frac{\tilde{m}_{1,t}}{1 - \omega_t} - \eta. \quad (\text{H.47})$$

Finally, when both agents hold positive portfolios, the optimal portfolios are $\vec{w}_t^R = (\sigma_t \sigma_t')^{-1} (\vec{\mu}_t - r_t \mathbf{1}_{2 \times 1})$, $\vec{w}_t^I = (\sigma_t \sigma_t')^{-1} (\vec{\mu}_t - r_t \mathbf{1}_{2 \times 1} + \sigma_{1,t} \vec{\eta})$. Repeating the arguments in equations (H.37)–(H.42), we obtain $\kappa_{1,t} = \tilde{m}_{1,t} \sigma_{1,t} - (1 - \omega_t) \eta$. ■

Proof of Proposition 7. It remains to derive the differential equation in Proposition 7. Using the market clearing condition $\sum_{i \in \{I, R\}} \omega_t^i w_{1,t}^i = \tilde{m}$, and applying Ito's Lemma to $\omega_t^i = \frac{W_t^i}{W_t^I + W_t^R}$ leads to

$$d\omega_t^i = \mu_{\omega,t}^i dt + \sigma_{\omega,t}^i dB_{1,t} \quad (\text{H.48})$$

with

$$\begin{aligned} \mu_{\omega,t}^i &= \omega_t^i [(w_{1,t}^i - \tilde{m}) \sigma_{1,t} (\kappa_t - \sigma_{1,t} \tilde{m}) + w_{1,t}^i \lambda_t^i + \tilde{n}_t] + \theta (\nu_t^i - \omega_t^i), \\ \sigma_{\omega,t}^i &= \omega_t^i (w_{1,t}^i - \tilde{m}) \sigma_{1,t}, \end{aligned}$$

and⁶⁶

$$\tilde{n}_t \equiv - \sum_{i \in \{I, R\}} w_{1,t}^i \omega_t^i \lambda_t^i = \frac{y_t \tilde{m}}{1 - y_t} f_t (1 - \tau).$$

Since $\frac{\phi_1}{\phi_2} \approx 0$, the aggregate endowment follows a geometric Brownian motion in the limit, and the interest rate is constant $r_t = r$. Accordingly, the price of a stock of type 1 follows the dynamics

$$\frac{dP_{1,t,s} + D_{1,t,s} dt}{P_{1,t,s}} = (r + \kappa_{1,t} \sigma_{1,t}) dt + \sigma_t dB_{1,t}. \quad (\text{H.49})$$

Applying Ito's Lemma to the product $P_{1,t,s} = p(\omega_t) D_{1,t,s}$ also implies that

$$\frac{dP_{1,t,s}}{P_{1,t,s}} = \frac{dp_t}{p_t} + \frac{dD_{1,t,s}}{D_{1,t,s}} + \frac{p'(\omega_t)}{p(\omega_t)} \sigma_{\omega,t}^R \sigma_{1,D} dt. \quad (\text{H.50})$$

66. Using $\sum_{i \in \{I, R\}} w_{1,t}^i \omega_t^i = \tilde{m}_t$, the definition $y_t = -\frac{w_{1,t}^R \omega_t^I \mathbf{1}_{\{w_{1,t}^R < 0\}}}{w_{1,t}^I \omega_t^I}$ and the definition of λ_t^i leads to

$$- \sum_{i \in \{I, R\}} w_{1,t}^i \omega_t^i \lambda_t^i = \frac{y_t \tilde{m}}{1 - y_t} f_t (1 - \tau).$$

Combining (H.49) with (H.50) and using $\sigma_{1,t} = \frac{p'(\omega_t)}{p(\omega_t)}\sigma_{\omega,t}^R + \sigma_{1,D}$ and Ito's Lemma to compute the drift of $\frac{dp_t}{p_t}$ leads to

$$\frac{1}{2} \frac{\partial^2 p}{\partial \omega_t^2} (\sigma_{\omega,t}^R)^2 + \frac{\partial p}{\partial \omega_t} (\mu_{\omega,t}^R + \sigma_{\omega,t}^R \sigma_{1,D}) - p \times (r + \delta_1 + \kappa_{1,t} \sigma_{1,t}) + 1 = 0, \quad (\text{H.51})$$

which in turn leads to (F.6) after substituting $\sigma_{1,t} = \frac{p'(\omega_t)}{p(\omega_t)}\sigma_{\omega,t}^R + \sigma_{1,D}$. ■

Proof of Proposition 8. Using a standard “guess-and-verify” approach, the modified first-order conditions for portfolio choice of the two investors (conditional on the shorting market being active, i.e., $w^R < 0$) give

$$w^R = \frac{\mu - r + f + \chi}{\sigma^2} = \frac{\kappa}{\sigma} + \frac{f + \chi}{\sigma^2} \quad (\text{H.52})$$

$$w^I = \frac{\mu - r + \tau f y}{\sigma^2} = \frac{\kappa + \eta}{\sigma} + \frac{\tau f y}{\sigma^2}. \quad (\text{H.53})$$

The consumption-to-wealth ratio for all investors continues to be $\rho + \pi$ and therefore the price-dividend ratio is constant and given by $\frac{1}{\rho + \pi}$. As a result, the volatility σ is constant. From this point on, repeating the steps summarized by equations (15)–(20), we obtain

$$y = \frac{\eta - \frac{\chi}{\sigma} - \frac{\sigma}{1-\omega} - \frac{f}{\sigma}(1 - \tau y)}{\eta - \frac{\chi}{\sigma} + \frac{\sigma}{\omega} - \frac{f}{\sigma}(1 - \tau y)}. \quad (\text{H.54})$$

We note that (H.54) is the same as (20), but with η replaced by $\eta - \frac{\chi}{\sigma}$. Applying the implicit function theorem to (H.54) and focusing on the maximum shorting equilibrium (which is the “stable” equilibrium) shows that $\frac{dy}{d\chi} < 0$. In addition, equation (H.52) together with $w_t^R < 0$ and (29) implies $\frac{d\sigma_\omega}{d\chi} > 0$.⁶⁷ These observations prove parts (i) and (ii).

For part (iii), consider the derivative of $\sigma_y^2 = \left(\frac{dy}{d\omega}\sigma_\omega\right)^2$:

$$\frac{d\sigma_y^2}{d\chi} = 2 \left(\frac{dy}{d\omega}\right)^2 \sigma_\omega^2 \left(\frac{\frac{d\sigma_\omega}{d\chi}}{\sigma_\omega} + \frac{\frac{d^2 y}{d\omega d\chi}}{\frac{dy}{d\omega}}\right). \quad (\text{H.55})$$

The term $\frac{1}{\sigma_\omega} \frac{d\sigma_\omega}{d\chi}$ is negative because $\frac{d\sigma_\omega}{d\chi} > 0$ and $\sigma_\omega < 0$. The remainder of the proof shows that also the second term inside the last parentheses, $\left(\frac{dy}{d\omega}\right)^{-1} \frac{d^2 y}{d\omega d\chi}$, is negative for small ω . To

67. Using the expression for the equilibrium Sharpe ratio gives

$$\begin{aligned} w^R - 1 &= \frac{\kappa}{\sigma} + \frac{f + \chi}{\sigma^2} - 1 = -(1 - \omega) \left(\frac{\eta}{\sigma} - \frac{\chi}{\sigma^2}\right) + \frac{f}{\sigma^2} (1 - \omega) (1 - \tau y) \\ &= -(1 - \omega) \left[\left(\frac{\eta}{\sigma} - \frac{\chi}{\sigma^2}\right) - \frac{f}{\sigma^2} (1 - \tau y)\right] < 0. \end{aligned}$$

Therefore $\frac{d(w^R - 1)}{d\chi} = (1 - \omega) \left[\frac{1}{\sigma^2} - \frac{f}{\sigma^2} \tau \frac{dy}{d\chi}\right] > 0$, since $\frac{dy}{d\chi} < 0$.

start, the implicit function theorem gives the following expression⁶⁸ for $\frac{dy}{d\omega}$:

$$\frac{dy}{d\omega} = \sigma \frac{\frac{(1-\omega)^2 y - \omega^2}{(1-\omega)^2 \omega^2}}{\eta + \frac{\sigma}{\omega} - \frac{f}{\sigma}(1 - \tau y) - \frac{\chi}{\sigma} - \tau(1 - y) \frac{f}{\sigma}}. \quad (\text{H.56})$$

We next show that $\frac{dy}{d\omega} > 0$ for small enough ω . To see this, note that the larger root of equation (H.54) satisfies $\lim_{\omega \rightarrow 0} y(\omega) = 0$. Moreover, re writing equation (H.54) as

$$y \left(\eta - \frac{\chi}{\sigma} + \frac{\sigma}{\omega} - \frac{f}{\sigma}(1 - \tau y) \right) = \eta - \frac{\chi}{\sigma} - \frac{\sigma}{1 - \omega} - \frac{f}{\sigma}(1 - \tau y)$$

and taking the limit as $\omega \rightarrow 0$ on both sides implies that

$$\sigma \lim_{\omega \rightarrow 0} \frac{y}{\omega} = \lim_{\omega \rightarrow 0} y \left(\eta - \frac{\chi}{\sigma} + \frac{\sigma}{\omega} - \frac{f}{\sigma}(1 - \tau y) \right) = \lim_{\omega \rightarrow 0} \eta - \frac{\chi}{\sigma} - \frac{\sigma}{1 - \omega} - \frac{f}{\sigma} > 0. \quad (\text{H.57})$$

Therefore, for small enough ω $\frac{y}{\omega^2}$ is arbitrarily large and therefore the numerator in (H.56) is positive. We also note that the denominator of $\frac{dy}{d\omega}$ is also positive for the (stable) equilibrium associated with y^+ . Accordingly, $\frac{dy}{d\omega} > 0$ for small enough ω .

Next, we totally differentiate $\frac{dy}{d\omega}$ with respect to χ and observe that the sign of $\left(\frac{dy}{d\omega}\right)^{-1} \frac{d^2 y}{d\omega d\chi}$ is the same as the sign of

$$\begin{aligned} & \left(\frac{(1-\omega)^2}{(1-\omega)^2 y - \omega^2} + \frac{2\tau \frac{f}{\sigma}}{\eta + \frac{\sigma}{\omega} - \frac{f}{\sigma}(1 - \tau y) - \frac{\chi}{\sigma} - \tau(1 - y) \frac{f}{\sigma}} \right) \frac{dy}{d\chi} \\ & - \frac{-\frac{1}{\sigma}}{\eta + \frac{\sigma}{\omega} - \frac{f}{\sigma}(1 - \tau y) - \frac{\chi}{\sigma} - \tau(1 - y) \frac{f}{\sigma}}. \end{aligned} \quad (\text{H.58})$$

As $\omega \rightarrow 0$, the term $\eta + \frac{\sigma}{\omega} - \frac{f}{\sigma}(1 - \tau y) - \frac{\chi}{\sigma} - \tau(1 - y) \frac{f}{\sigma}$ approaches infinity, and hence the sign of the expression (H.58) is the same as the sign of $\frac{(1-\omega)^2}{(1-\omega)^2 y - \omega^2} \frac{dy}{d\chi}$, which as we argued above has the same sign as $\frac{dy}{d\chi}$. An application of the implicit function theorem to (H.54) shows that $\frac{dy}{d\chi} < 0$ (for the stable equilibrium y^+). Accordingly, $\frac{d\sigma_y^2}{d\chi} < 0$ for sufficiently small ω . ■

68. Note that equation (H.54) can be written as

$$H(y) \equiv y \left(\eta + \frac{\sigma}{\omega} - \frac{f}{\sigma}(1 - \tau y) - \frac{\chi}{\sigma} \right) - \left(\eta - \frac{\sigma}{1 - \omega} - \frac{f}{\sigma}(1 - \tau y) - \frac{\chi}{\sigma} \right) = 0$$

and therefore (H.56) follows from the implicit function theorem, $\frac{dy}{d\omega} = -\frac{H_\omega}{H_y}$.

Table I.1: Summary Statistics of Shorting fees.

Size quintile	Percentile				
	50 th	75 th	90 th	95 th	99 th
1	0.41%	0.84%	2.96%	7.43%	28.48%
2	0.38%	0.50%	1.44%	3.97%	19.38%
3	0.36%	0.43%	0.86%	2.04%	12.30%
4	0.34%	0.39%	0.53%	1.02%	7.39%
5	0.35%	0.38%	0.43%	0.50%	1.72%
Total	0.37%	0.51%	1.24%	2.99%	13.85%

Lending fees by stock market capitalization quintile. Each year, we form 5 portfolios of Russell 3000 constituents sorted into size quintiles based on end-of-prior-year market capitalization. Within each size quintile, we compute the p^{th} percentile, $p \in \{50, 75, 90, 95, 99\}$, of daily shorting fees over the following year. We then report the time-series average of these percentiles from 2006 to 2021. Daily shorting fees from 2006 to 2021 are from Markit and are reported as annualized percentage rates.

I Additional Empirical Results

I.1 Summary Statistics — IHS Markit

We start by reporting some summary statistics on lending fees. In Table I.1, we group Russell 3000 constituents based on their end-of-prior-year market capitalization into five quintiles. We then fix the set of stocks in each quintile over the subsequent year and compute various statistics (median, 75th percentile, etc.) of the daily lending fees for the stocks in each quintile. We then average across the years. The table shows that the median lending fee ranges between 0.35% and 0.41%. However, the table also shows that some of the observations on lending fees can become quite large. For instance, for stocks that are in the size portfolios 1, 2, and 3, the 95-th percentile of fees exceeds 2% and the 99-th percentile exceeds 7% for stocks in portfolios 1, 2, 3, and 4. This table suggests that sometimes even relatively large stocks (by market capitalization) can exhibit sizable lending fees.

Table I.2 helps to illustrate this last point in greater detail. Specifically, Table I.2 reports some stock-level statistics on lending fees, and in particular the fraction of Russell 3000 constituents for which a given percentile of shorting fees across time exceeds certain cutoffs. The table shows that 96% of Russell 3000 constituents exhibit a lending fee in excess of 1% at some point between 2006 and 2021, while 45% of stocks exhibit a fee in excess of 5% at some point over that same time period. But even if we leave these extreme observations aside, and focus on — say — the 95-th percentile of the distribution of lending fees at the stock level, the numbers are large: 31% of Russell 3000 constituents exhibit a lending fee in excess of one percent for 5 out of 100 trading days, while 18% of Russell constituents exhibit lending fees in excess of 3% for 5 out of 100 trading days.

Table I.2: Stock-level distribution of shorting fees.

Percentile	Shorting fee cutoffs				
	$\geq 1\%$	$\geq 2\%$	$\geq 3\%$	$\geq 5\%$	$\geq 10\%$
90 th	0.23	0.16	0.12	0.09	0.05
95 th	0.31	0.21	0.18	0.14	0.07
99 th	0.50	0.32	0.26	0.21	0.13
99.5 th	0.62	0.38	0.30	0.23	0.14
Maximum	0.96	0.79	0.66	0.45	0.27

Fraction of Russell 3000 constituents for which the indicated percentile (first column) of daily shorting fees exceeds the cutoff noted in the header row. For example, the bottom rightmost number (0.27) means that 27% of the stocks in the Russell 3000 had a maximum daily shorting fee in excess of 10%. Similarly, the number 0.12 in the top row/ middle column indicates that 12% of the stocks have a lending fee in excess of 3 percent for one out of the ten trading days. Daily shorting fees from 2006 to 2021 are from Markit and are reported as annualized percentage rates.

I.2 Heterogeneous $h'(y)$

For our baseline results we pooled observations across all stocks and estimated a single function $h'(y)$. Figure I.1 shows results for the case where we allow $h'(y)$ to differ for each stock. Specifically, we focus on observations that are on special (DCBS > 1) and estimate a separate $h'(y)$ for each Russell 3000 constituent. We then evaluate $h'(y)$ for different values of y for each stock separately. Subsequently, we pool all $h'(y)$ values across all stocks and present them as a bin-scatter diagram.⁶⁹ Since stock-level estimates of $h'(y)$ are noisy, we trim stock-level estimates of $h'(y)$ at the 5th and 95th percentile levels. (Results are similar if we don't trim and instead report medians by shorting-utilization bin.) The main conclusion from Figure I.1 is similar to our conclusion in the text: for low values of y , $h'(y)$ is negative.

69. Since the observations per stock are not in the millions (as they are for the pooled regressions in the text), it is computationally feasible to use a kernel regression estimator with automatic, cross-validated, bandwidth selection. We present the results for this alternative estimation method, as a check that our conclusions are not driven by whether we use kernels or splines to estimate the non-parametric regression.

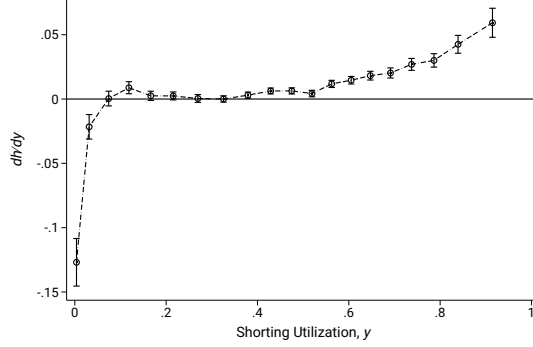


Figure I.1: $h'(y)$, binned-means from stock-level estimates using local-linear non-parametric kernel regressions. For each stock, we estimate the marginal effect $h'(y)$ at 11 points, corresponding to the stock-level deciles, as well as the 5th and 95th percentiles of utilization for observations exhibiting a Daily Cost of Borrowing Score (DCBS) over 1. Error bars represent 95% confidence intervals around bin means. Data on shorting fees and shorting utilization are from Markit over the period 2006 to 2021.

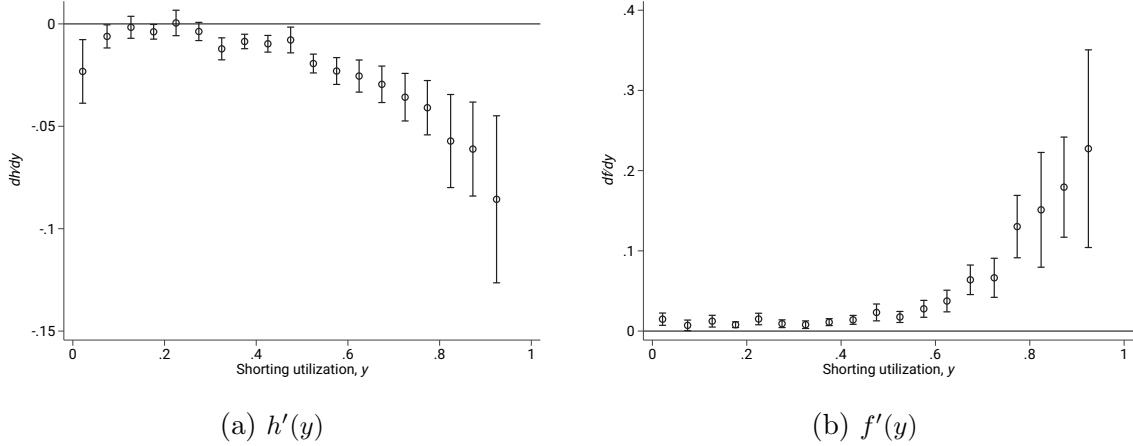


Figure I.2: $h'(y)$ and $f'(y)$, binned-means from jump discontinuities. Estimated marginal effects are calculated by dividing observed weekly changes in $h(y) = f(y)(1 - \tau y)$ by observed weekly changes in shorting utilization, restricting attention to weeks in which the magnitude of the change in shorting utilization magnitude exceeds 10.5%, four times the standard deviation of absolute changes in utilization. We calibrate τ to be 0.8 based on industry literature on the pass-through of shorting fees to institutional investors. Sample consists of daily observations of shorting fees and shorting utilization for Russell 3000 constituents. Error bars represent 95% confidence intervals around bin means. Data on shorting fees and shorting utilization are from Markit over the period 2006 to 2021.

Table I.3: Determinants of Utilization Jump Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$1_{\text{Fail - Satisfy}}$	-5.14*** (-5.63)	-4.78*** (-5.67)	-3.62*** (-4.30)	-5.28*** (-5.79)	-5.00*** (-5.51)	-3.85*** (-4.37)	-2.96*** (-3.82)
Shorting Fee		59.68*** (11.03)					52.64*** (9.65)
Size Quintile							
2			-8.86*** (-14.79)				-7.10*** (-12.98)
3			-10.76*** (-14.33)				-8.17*** (-11.38)
4			-13.17*** (-14.62)				-8.35*** (-9.35)
5			-12.62*** (-9.53)				-8.88*** (-6.60)
Var of Returns				63.37** (1.99)			24.73** (2.31)
Turnover				36.80** (1.99)			-4.03 (-0.24)
Debt/Total Assets					-4.61*** (-3.36)		-1.36 (-1.16)
Log Book/Market					0.86*** (2.72)		1.22*** (3.86)
1_{Option}						-9.29*** (-11.82)	-5.80*** (-8.27)
1_{NASDAQ}						1.36* (1.88)	-0.04 (-0.05)
N	1975	1975	1975	1975	1975	1975	1975

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

This table reproduces the estimates of Table 1 and, in addition, reports the coefficients on the various stock-level characteristics. All columns also include dummies for the quintile that each stock belongs to based on the t-statistic of satisfying the assumption of Proposition 4 (see the description in the text). For ease of comparison with Table 1 we report (in the first row) only the coefficient on the dummy variable that compares stocks in the fifth quintile vs. the first (base) quintile under the heading $1_{\text{Fail-Satisfy}}$.

J Details and Additional Results for Section 7

J.1 Measuring ticker discussion on WallstreetBets

Our measure of ticker mentions on WallstreetBets is constructed as follows. We use the PushshiftAPI to collect all submissions posted on WallstreetBets subreddit from January 1, 2020 through February 7, 2021 (Baumgartner et al. 2020b). For each submission, we observe the title text, the body of the submission, the author of the submission, and the time of the submission.

We then identify all cases in which these tickers are mentioned in submissions, irrespective of whether they are prefixed with a dollar sign. To address the possibility of falsely identifying tickers, we require that, if the ticker is a common word in the written English language, it must be prefaced by a dollar sign. For example, AT&T’s ticker T is also a common word in written English, and thus we require that the text “\$T” appear in a submission for it to be considered as mentioned AT&T. We consider a ticker as being mentioned in a submission if it appears in either the title or the body of the submission. We identify common word-stems based on the Google Trillion Word Corpus (Michel et al. 2011b). In a robustness check, we account for the downward bias this restriction introduces by scaling common-word tickers by an in-sample estimated adjustment factor. This adjustment leaves the relative ranking of ticker mentions largely unchanged. We estimate the adjustment factor by comparing the frequency of tagged ticker mentions versus un-tagged ticker mentions for the set of tickers which do not commonly appear in written English.

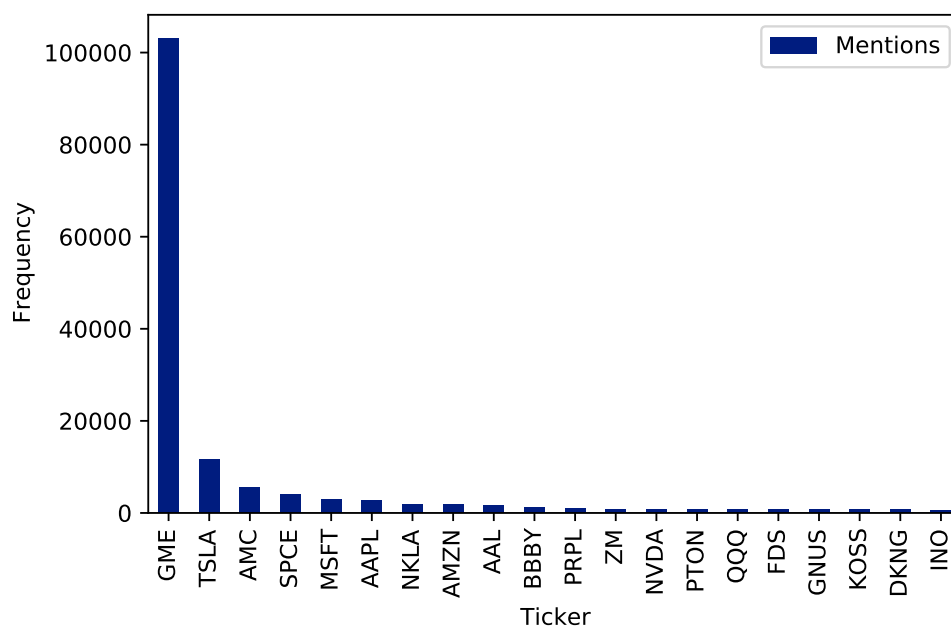
Revised submissions and comments. Authors of Reddit comments have the ability to edit their comments even after the comment has been posted. The PushshiftAPI records the comment text as of a certain day, and does not update to reflect potential revised comments. The same constraint applies to the content body of submissions. Titles of submissions cannot be revised and thus do not have this measurement problem.

Missed tickers Tickers that, for whatever reason, are never tagged with a leading dollar sign will be omitted from our dataset. Similarly, we under-count the occurrences of tickers that are common words, owing to requiring they appear with a leading “\$” We attempt to correct for this by scaling the observed counts for common word tickers. For AAPL and GME, which are not common word tickers, the ticker appears with the leading “\$” roughly 20% of the time. We can thus simply multiply our observed frequencies by a factor of five to adjust for the more stringent matching procedure. As can be seen in Figures J.1a and J.1b, the adjustment does not have a significant impact on the relative popularity of the top tickers.

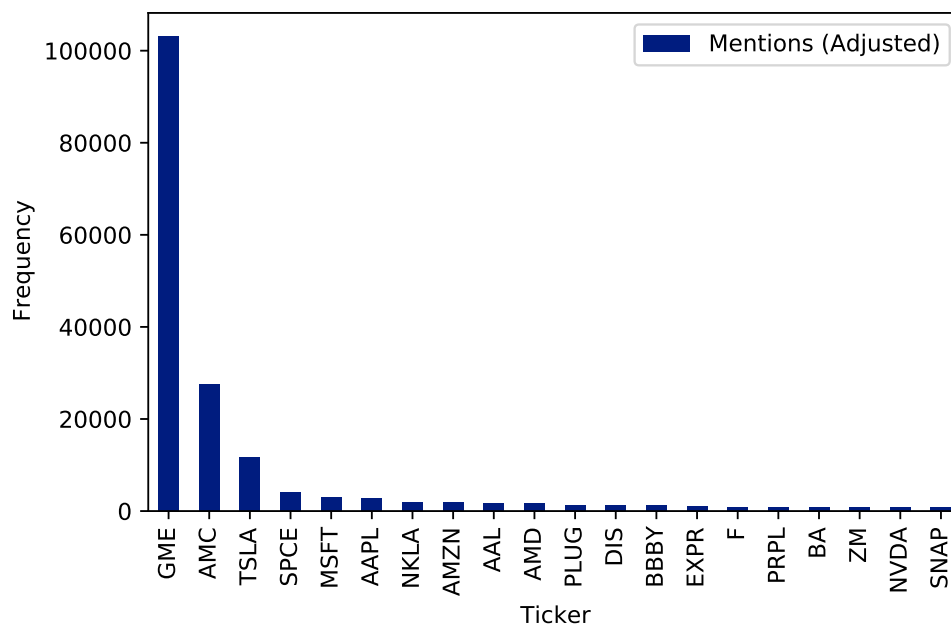
In some cases, users may choose to refer to the company by its name, rather than by its ticker. We do not attempt to identify mentions of companies by name.

J.2 Measuring retail trading

We adopt the methodology of Boehmer et al. (2021) to identify retail trades in the TAQ data. We briefly summarize the methodology here and refer readers to the paper for details.



(a) Submissions mentioning each Ticker



(b) Submissions mentioning each Ticker, adjusted for word-ticker overlap

Figure J.1: Popular Tickers on WallstreetBets (January 1, 2020 – February 7, 2021).

The intuition behind the methodology is the knowledge that retail trades are often executed by wholesalers or via broker internalization, rather than on the major trading exchanges. These trades appear in the TAQ consolidated tape data under the exchange code “D.” These trades are given a small price improvement on the order of tenths of a penny as a means to induce brokers to route orders to the wholesaler. Similarly, brokers which internalize retail trades offer a subpenny price improvement in order to comply with Regulation 606T. Importantly, institutional trades are rarely, if ever, internalized or directed to wholesalers and their trades are usually in round penny prices, with the notable exception of midpoint trades.

The methodology of Boehmer et al. (2021) uses these institutional details to identify retail trades in the TAQ consolidated tape data. Trades flagged with exchange code “D” and with a subpenny amount in the set $(0, 0.40) \cup (0.60, 1.00)$ are identified as retail trades. Splitting these trades further, retail trades with subpenny amounts between zero- and forty-hundredths of a penny are labeled as “sell orders,” whereas subpenny amounts between sixty- and one hundred-hundredths are considered “buy orders.” The midpoint trades are excluded to avoid mis-classifying institutional trades executed at midpoints as retail trades.

J.2.1 Challenges

Derivatives The TAQ data only contains trades of equities. Options offer another way to benefit for investors to benefit from increases in the price of stock. As an added advantage for retail investors, options offer embedded leverage greater than what might otherwise be available through their broker. The Boehmer et al. (2021) methodology relies on institutional details to identify off-exchange retail trades, and thus cannot reliably identify replication trades by market makers. See Barber et al. (2024) for additional discussion of the limitations of the aforementioned methodology.

J.3 Betting against the shorts portfolio

As is standard in the literature, we restrict attention to common shares of COMPUSTAT firms which trade on the NYSE, AMEX, and NASDAQ exchanges. We further exclude companies for whom no share class has a price exceeding \$1. The strategy equally weights each firm in the top decile, shorts the market index, and reconstitutes 8 trading days following the disclosure date, which is the first opportunity following the public dissemination of the short interest data.

	Highly Shorted Stocks	Excl. Popular Reddit Stocks	Excl. Small Stocks
<i>Panel A: November 2020</i>			
r^{EW}	0.164 (4.330)	0.161 (4.249)	0.226 (5.077)
r^{VW}	0.093 (3.208)	0.092 (3.186)	0.132 (3.446)
r_{FF3}^{EW}	0.086 (3.641)	0.083 (3.512)	0.157 (4.373)
r_{FF3}^{VW}	0.044 (1.871)	0.042 (1.819)	0.080 (2.461)
<i>Panel B: December 2020</i>			
r^{EW}	0.057 (1.509)	0.061 (1.606)	0.019 (0.420)
r^{VW}	0.033 (1.128)	0.036 (1.243)	0.020 (0.517)
r_{FF3}^{EW}	0.017 (0.727)	0.021 (0.882)	-0.000 (-0.009)
r_{FF3}^{VW}	0.013 (0.546)	0.015 (0.671)	0.008 (0.257)
<i>Panel C: January 2021</i>			
r^{EW}	0.271 (6.835)	0.233 (5.865)	0.156 (3.576)
r^{VW}	0.194 (6.658)	0.160 (5.576)	0.182 (4.762)
r_{FF3}^{EW}	0.205 (8.685)	0.167 (7.059)	0.113 (3.135)
r_{FF3}^{VW}	0.163 (6.895)	0.128 (5.598)	0.155 (4.741)

Table J.1: Portfolio returns (November 2020–January 2021). Test of whether the monthly return to the strategy of betting against the shorts is “abnormal” in November 2020 (Panel A), December 2020 (Panel B), and January 2021 (Panel C). The table reports the coefficient and the t -statistic of the month dummy variable that takes the value of one for the month listed in the title of the panel and zero otherwise from the regression:

$$r_{\text{Betting against the shorts}} = \text{const.} + \text{month dummy} + \beta' F_t + \varepsilon_t.$$

The first two rows of each panel do not control for any factor exposures and refer to equal-weighted (EW) and value-weighted (VW) returns, respectively. The last two rows of each panel control for Fama-French 3-factor exposures.

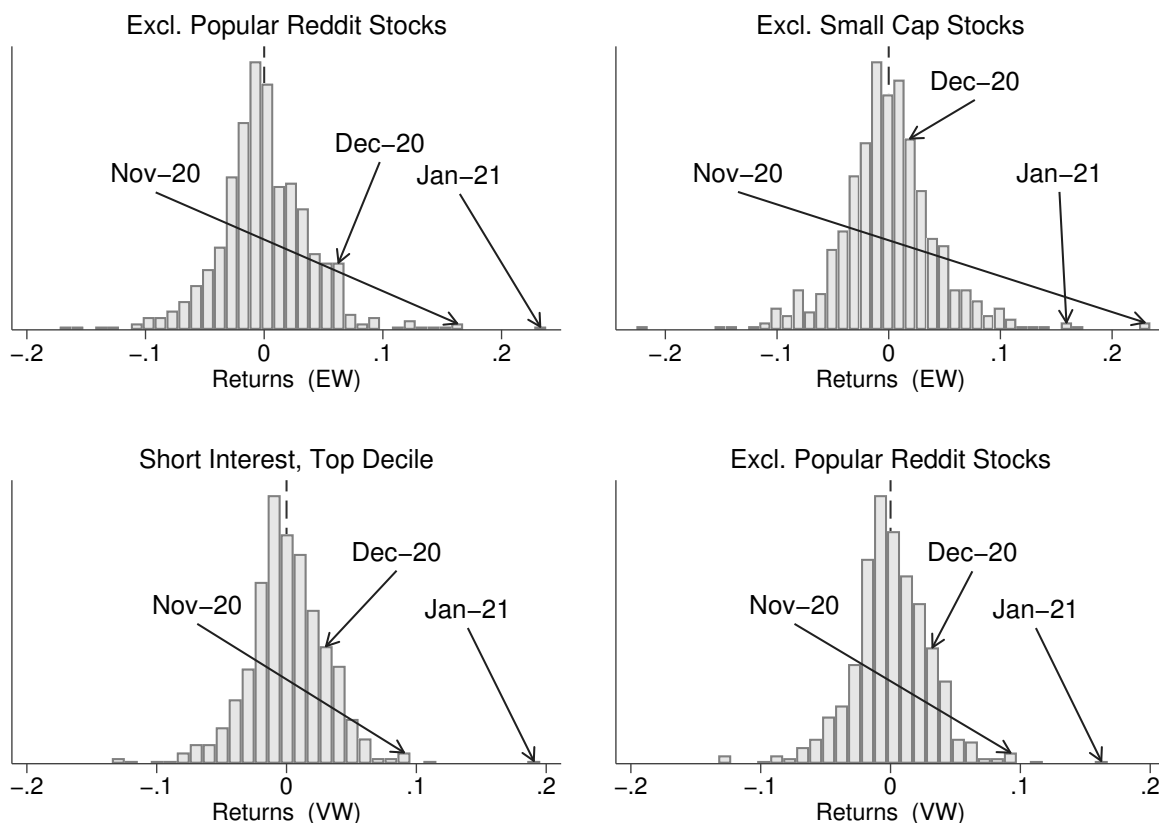


Figure J.2: Monthly returns (1973–2021). Histograms show monthly returns to a trading strategy long stocks in the top decile of short interest and short the market index. The top-left plot depicts equal-weighted returns, excluding the six most-popular stocks discussed on Reddit (AMC, BBBY, GME, SPCE, TLRY, and TSLA). The top-right plot depicts equal-weighted returns, further excluding small market capitalization stocks. The bottom-left plot depicts value-weighted returns. The bottom-right plot depicts value-weighted returns, excluding popular stocks discussed on Reddit. The arrows indicate the portfolio returns in the months of November and December 2020 and January 2021.

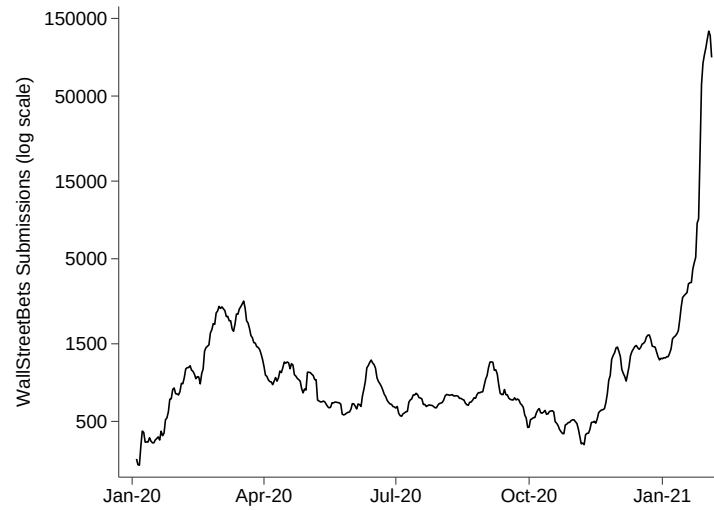


Figure J.3: Seven-day moving average of daily submissions to the WallstreetBets subreddit (January 1, 2020 – February 7, 2021). The vertical axis is on a logarithmic scale.

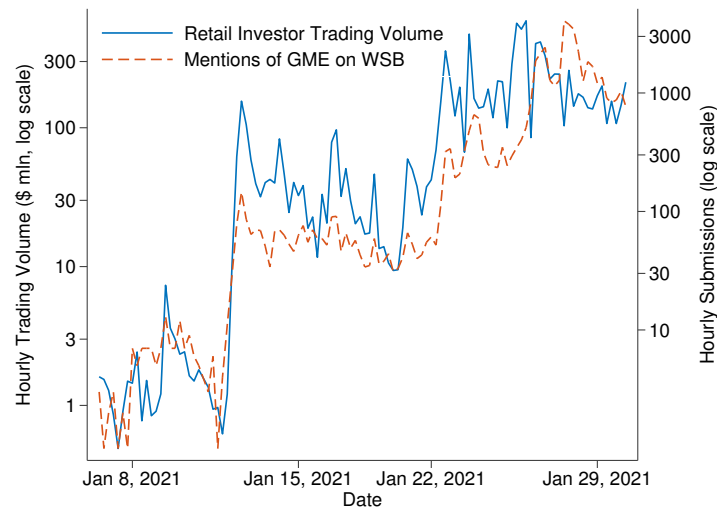


Figure J.4: Retail trading volume in GME (January 7 – January 29, 2021). Hourly trading volume in GME, measured using the methodology of Boehmer et al. (2021), plotted together with hourly mentions of the GME ticker on the WallStreetBets subreddit. Both vertical axes are on logarithmic scales.

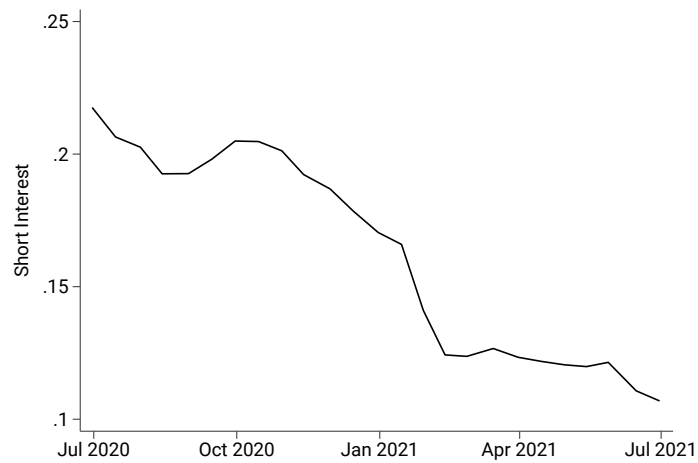


Figure J.5: Aggregate short interest (July 2020–June 2021). The figure plots value-weighted short interest for highly shorted stocks as of October 31, 2020. Highly shorted stocks are defined as the stocks in the top decile of the Russell 3000, ranked by short interest. The identities of these stocks is fixed and their short interest is plotted over the preceding four and subsequent eight months.

References - Online Appendix

- Barber, Brad M., Xing Huang, Philippe Jorion, Terrance Odean, and Christopher Schwarz. 2024. “A (Sub)penny for Your Thoughts: Tracking Retail Investor Activity in TAQ.” *The Journal of Finance*.
- Baumgartner, Jason, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. *The Pushshift Reddit Dataset*.
- Boehmer, Ekkehart, Charles M. Jones, Xiaoyan Zhang, and Xinran Zhang. 2021. “Tracking Retail Investor Activity.” *The Journal of Finance* 76 (5): 2249–2305.
- Karatzas, Ioannis, and Steven Shreve. 2012. *Brownian motion and stochastic calculus*. Vol. 113. Springer Science & Business Media.
- Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, Joseph P. Pickett, Dale Hoiberg, et al. 2011b. “Quantitative Analysis of Culture Using Millions of Digitized Books.” *Science* 331 (6014): 176–182.