The Participation Puzzle*

Federico Mainardi † Roxana Mihet‡ Laura Veldkamp §

March 25, 2026

Abstract

Financial technology was supposed to democratize finance. Despite declines in onboarding, trading, and data costs, participation in risky asset markets has fallen. We document the puzzle, offer a resolution, test the explanation and quantify it. We argue that when all participants can access data more cheaply, wealthier investors, with their larger portfolios, acquire relatively more data than others. Their informed trading reduces expected returns for less informed traders, lowering the marginal household's gains from risky investing. The quantitative model shows that lower entry costs are not sufficient to overcome the adverse selection cost. Consistent with the model predictions, we find that returns and return premia accrue disproportionately to wealthier and more sophisticated investors. Finally, because informed investors both exhibit more skill and hold riskier portfolios, the framework reconciles competing explanations for financial income inequality.

Keywords: Participation, equity market, inequality, Gini coefficient, information frictions, wealth distribution, risk aversion.

JEL-Codes: E21, G11, G14, L1, L15.

*This paper supersedes "Financial Technology and Inequality." The authors thank ... for insightful suggestions, as well as participants at numerous conferences and seminars for useful feedback. Roxana Mihet acknowledges generous funding for this project from The Sandoz Family Foundation - Monique de Meuron Programme. All authors declare no conflicts of interest related to this project. All errors are solely the authors' responsibility.

†Federico Mainardi. Address: Columbia Business School, 665 W 130th St, New York, NY 10027, United States. Email: fm2913@columbia.edu.

‡Roxana Mihet. Address: Swiss Finance Institute at the University of Lausanne, Quartier UNIL-Chamberonne, Extranef 138, Lausanne, 1015 Vaud, Switzerland. Email: roxana.mihet@unil.ch.

§Laura Veldkamp. Address: Columbia Business School, 665 W 130th St, New York, NY 10027, United States. Email: lv2405@columbia.edu.

The modern era of financial innovation was heralded as the democratization of finance, promising a future where technological progress would lower barriers to entry and narrow the wealth gap. During the 1990s, this promise seemed to be materializing: as digital brokerage platforms emerged and trading fees plummeted, household stock market participation surged. Yet, in the early 2000s, this progress hit a wall. Despite a period of even more massive financial innovation, household participation in risky asset markets has since stalled or even declined (Gomes, Haliassos and Ramadorai, 2021).[1] This participation puzzle is particularly striking because it coincides with a period where the tools to process financial data, and produce sophisticated analytics have become exponentially cheaper and more widely available. If the costs of both access and information have fallen, why has the democratization of the stock market stopped in its tracks?

We distinguish between two distinct technological innovations: trading technology, which facilitates entry and onboarding, and data provision, which is often bundled with platforms, but offers sophisticated market signals and analytics. Empirical proxies for trading costs, such as the minimum fees for passive, diversified mutual funds, declined rapidly during the 1990s but have remained virtually flat for the past two decades. In contrast, the costs of data provision have continued to decline. While the flattening of entry costs explains why participation stopped growing (Favilukis, 2013), it does not explain why it has retrenched. Our finding is that falling data costs, rather than inviting more households into the market, may be the very force pushing them out.

To explore this mechanism, we develop a micro-founded, three-layered demand system where equilibrium price effects determine participation, information, and investment demand jointly. In this framework, investors first choose whether to pay a fixed entry cost to access the market. Second, they decide whether to become informed and

[1] Using SCF data, Poterba (2018) shows that stock ownership rose sharply into the early 2000s, but that both overall stock ownership and (more starkly) direct stock ownership declined thereafter. Wolff (2014, 2017) also documents the post-2005 decline in stock ownership.

how much data to acquire or process. Finally, they select risky asset investments.

The core of the puzzle is resolved through an adverse selection channel triggered by the increasing returns to scale inherent in financial data. Because wealthier investors take larger risky positions, they have a much higher incentive to utilize crashing Data Provision costs to acquire precise private signals. As the informed sector becomes more precise, equilibrium prices become more informationally efficient. However, this increased efficiency is a double-edged sword: it reduces the risk-adjusted returns available to uninformed middle-class participants. To avoid being picked off by more aggressively informed wealthy traders, these marginal households optimally exit the market, forgoing the equity premium and choosing to stay in riskless assets.

Does the ability to pool resources into funds allow less-wealthy households to share data costs and overcome informational disadvantages? Not necessarily. Empirically, access to the most attractive delegated vehicles is often tiered by account size: lower-fee share classes and higher-quality intermediaries typically require substantially larger minimum investments, limiting access for middle-class households. Theoretically, a model extension shows that delegation does not eliminate the participation paradox because investors must still search for funds, and the wealthy have greater incentives to seek out and profit from the best ones. Thus, pooling changes how information is purchased, but not the equilibrium force that links cheaper information to lower participation among less-informed households.

The model delivers five testable hypotheses. First, participation should rise with wealth. While that prediction is common to many theories, it shows that ours also speaks to the well-known facts. Second, when data costs fall, participation should decline particularly in the middle of the wealth distribution, because poorer and wealthier investors are infra-marginal participants. Third, among participating investors, wealthier and informed investors should earn higher average returns as well as (fourth) higher risk-adjusted returns. Fifth, their portfolios should load more on risk factors, because

better information makes predicted returns more accurate, which induces more risk taking. To the econometrician that does not possess the same superior information, larger risky positions simply appear as higher risk exposure.

We find support for these predictions in two complementary datasets, the SCF and the Addepar portfolio data. Using SCF data, we show that participation rises strongly with wealth, while the post-2000 decline in risky-asset participation is concentrated in the middle quintiles, consistent with the model's participation mechanism. Using portfolio-level data from Addepar, we find that average returns and alphas rise with wealth, and that investors more likely to be informed because advised by sophisticated intermediaries earn higher returns and higher alphas. We also find that these investors exhibit greater market beta, consistent with the model's prediction that superior information makes portfolios look riskier in unconditional data. Taken together, these findings confirm that the benefits of the modern data revolution remain concentrated among those with the scale to exploit them, rather than encouraging new financial market entry.

Our prediction about greater portfolio risk among wealthier and informed investors also helps to reconcile the debate about whether wealthy households earn higher financial returns by taking on more risk or trading with more skill. In our framework, wealthy investors' information has two effects: it shifts their portfolios to raise their risk exposure, as in Calvet, Bach and Sodini (2019), and it generates information rents that appear as persistent within–asset-class skill differences, as in Fagereng, Guiso, Malacrino and Pistaferri (2020). Thus, seemingly conflicting theories of skill versus risk conclusions can arise from the same underlying force: technological change that widens *relative* informational advantages.

Related Literature This paper contributes to the intersection of household finance, equilibrium information acquisition, and the burgeoning literature on the distributional

consequences of financial technology. Our contribution is to reconcile the simultaneous decline in financial costs with the stagnation of market participation.

Theories and evidence about information choice show that wealthy investors acquire more information and thus earn higher returns, which thickens the right tail of the income and wealth distributions (Peress, 2004; Kacperczyk, Nosal and Stevens, 2018; Azarmsa, 2022). Easley, O'Hara and Yang (2016) explore a similar equilibrium effect, where more information reduces the benefits to others of becoming informed. Recent findings by Dessaint, Foucault and Frésard (2024) and Katona, Painter, Patatoukas and Zeng (2025) show that professional investors' use of big data or alternative data sources triggers such a return divergence effect. Among wealthy households, Frost, Gambacorta and Gambacorta (2022); Bretscher, Gomes, Sabbatucci and Tamoni (2022); Gambacorta, Gambacorta and Mihet (2023); Balloch, Mainardi, Oh and Vokata (2025) find empirical evidence of informational advantages from higher realized returns. What we add is a participation margin. Peress (2005) has a participation margin, but no choice of how much information to use. Our contribution is to show that falling data costs means more intensive data use by some, which triggers return divergence; this causes financial market participation to decline.

Research on stock market participation identifies fixed entry costs, comprising brokerage fees and the time required for financial education, as a primary reason why many households avoid risky assets (Vissing-Jørgensen, 2002a; Vissing-Jorgensen, 2002b; Gomes and Michaelides, 2005; Favilukis, 2013). While this literature focuses on how lowering trading costs facilitated entry until 2005, we address the participation reversal: the fact that participation rates later stalled or retrenched, despite continued technological progress.

While our focus is on participation, this obviously makes contact with the debate regarding the drivers of return heterogeneity across the wealth distribution. Empirical evidence from Norway suggests that higher returns for the wealthy are driven by

financial sophistication and skill (Fagereng et al., 2020), while evidence from Sweden attributes these gains primarily to exposure to systematic risk Calvet et al. (2019). We find that these conclusions are not contradictory. Wealthier households acquire superior data (the skill channel). Their more accurate return forecasts provide the confidence and economic incentive to take larger, more remunerative risky positions (the risk channel). This persistent heterogeneity in returns is a key ingredient for explaining the high concentration of wealth at the top of the distribution (Benhabib, Bisin and Luo, 2017; Smith, Zidar and Zwick, 2021).

The rest of the paper is organized as follows. Section 1 details the data we use and documents the participation puzzle. Section 2 presents a simple model that resolves the participation puzzle and provides testable predictions. Section 3 provides empirical evidence of the theory's predictions. Section 4 quantifies the model. Section 5 concludes.

1 Motivating Facts

We begin by documenting the empirical patterns that motivate the paper.

An important distinction is between trading technologies that lower the cost of entering risky asset markets and data costs of acquiring and processing information. Panel A of Figure 1 plots one proxy of access costs, defined as the minimum expense ratio charged by purely passive U.S. equity mutual funds' retail share classes. These costs declined rapidly at the beginning of the century, falling from about 20 basis points in 2000 to roughly 5 basis points by 2005. After that, the trend flattened considerably: between 2005 and 2020, the same measure declined only from about 5 basis points to roughly 2.5 basis points. In other words, the large decline in the cost of obtaining basic diversified equity exposure had largely run its course by the mid-2000s.

Data costs evolved very differently. According to the 2018 Technology Policy Institute report, the key costs for large-scale data processing continued to fall rapidly over

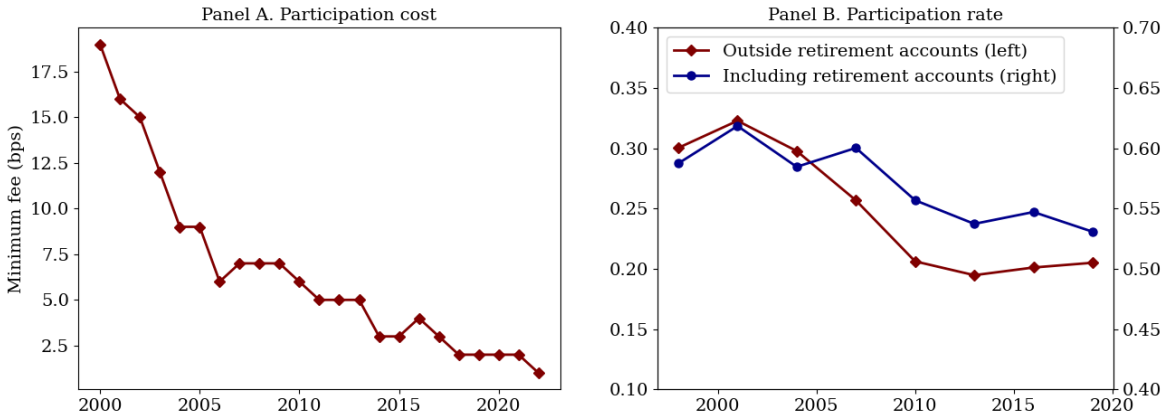


Figure 1: Access costs and participation in risky assets. Panel A plots our proxy for access costs. Panel B plots participation rates outside retirement accounts and including retirement accounts using SCF data.

the past two decades. For example, data transmission costs fell from about $1,200 per megabits per second in 1998 to about $0.02 in 2017. Over a shorter horizon, cloud processing prices fell by about one half between 2009 and 2016, while cloud storage prices fell by roughly 70 to 80 percent over the same period. Thus, while the decline in access costs had largely played out by 2005, the cost of producing, processing, and using financial data continued to fall steadily and substantially.

If the main decline in access costs had already occurred by the beginning of the century, while data costs continued to fall rapidly thereafter, one might expect participation in risky asset markets to keep rising. We now show that this is not what happened. Panel B of Figure 1 plots participation rates using data from the Survey of Consumer Finances (SCF). The first series measures participation outside retirement accounts and is defined as the fraction of households that hold risky financial assets outside defined-contribution plans. We define risky financial assets as financial assets other than cash-equivalent holdings, U.S. treasuries, and municipal bonds. In practice, this measure includes households with exposure to assets such as directly held equities, mutual funds, and corporate bonds.[2]

[2]Specifically, we define risky financial assets as the sum of stock mutual funds (*stmutf*), directly held stocks (*stocks*), other bond mutual funds (*obmutf*), combined mutual funds (*comutf*), other mutual

The second series measures participation inside and outside retirement accounts and is defined as the fraction of households that either hold any of those risky financial assets or have a position in a defined-contribution plan.[3] Focusing on participation outside retirement accounts is valuable because it reflects a household's active decision to enter risky markets, which maps more directly to the model's mechanisms than participation mediated or defaulted through retirement vehicles.

Despite the continued and rapid decline in data costs after 2005, participation in risky financial assets exhibits rapid retrenchment outside retirement accounts and moderately when accounting for risky assets held in retirement accounts. The combination of these facts is the central empirical challenge for the paper. Why did broad participation fail to expand when information became steadily cheaper to acquire and process? The goal of the paper is to provide a mechanism that reconciles these trends.

2 A Model Economy with Participation and Information Frictions

In order to explore the hypotheses that changes in information and trading technology can explain the decline in participation, we need a three-layered system. Investors need to make a participation choice, an information acquisition choice and a portfolio choice. Financial participation is also intimately linked to wealth. Therefore, we follow Peress (2004) in integrating wealth effects in an information choice model, by modeling wealth-dependent absolute risk aversion. Because we want to test predictions about informed and uninformed investors, we include both a linear and a convex cost to information. The positive linear (adoption) cost is what causes some investors to be uninformed. The convex (intensity) cost causes others to acquire more data when costs fall. Finally,

funds (*omutf*), and directly held bonds other than Treasuries and tax-exempt bonds (*obnd*).

[3] In the SCF, we measure defined-contribution wealth as the sum of *thrift* and *irakh*, which include assets held for instance through 401(k) plans and Individual Retirement Accounts (IRAs).

we need a market participation cost to explain non-participation.

This section formalizes these assumptions, explains why information and etnry costs have opposite effects on participation and then derives our testable predictions.

2.1 Environment

Agents. A unit mass of competitive investors $j \in [0,1]$ differ only in initial wealth $W_{0j} \in (0, \bar{W}_0]$, distributed according to V. Initial risky-asset holdings are normalized to zero; all initial wealth is held in the bond at $t = 0$.

Assets and payoffs. There is a risk-free bond in perfectly elastic supply, priced at 1 and paying gross return $r > 0$. There is one risky asset that expires and pays a fundamental dividend z at $t = 2$, with

$$z \sim \mathcal{N}(\mu_z, \sigma_z^2).$$

The risky asset is supplied by noise traders with (random) net supply

$$x = \mu_x + v, \qquad v \sim \mathcal{N}(0, \sigma_x^2),$$

independent of z and of all private signals.

Preferences. At the trading date, each investor j chooses risky holdings to maximize a mean–variance objective with increasing absolute risk tolerance τ_j:

$$\max_{q_j} \ \mathbb{E}_j[W_{2j}] - \frac{\tau_j^{-1}}{2}\mathbb{V}_j(W_{2j}), \qquad \text{where } \tau_j^{-1} \equiv \frac{W_{0j}}{\tau},\ \tau > 0. \tag{1}$$

Participation (entry) cost. To trade the risky asset, an investor must pay a fixed market-entry cost $F > 0$ at $t = 0$. Let $t_j \in \{0,1\}$ denote participation (i.e., trade in the risky asset). If $t_j = 0$, the investor holds only the bond and does not trade in the risky asset.

Private information and its cost. A participating investor may acquire a private

signal about z of the form

$$s_j = z + \varepsilon_j, \qquad \varepsilon_j \sim \mathcal{N}(0, d_j^{-1}),$$

where d_j is the chosen precision. Acquiring precision d_j costs

$$C(d_j) = c_1 d_j + \frac{c_0}{2} (d_j)^2, \qquad c_1 > 0,\ c_0 > 0. \tag{2}$$

The parameter c_1 captures an *adoption* component (a strictly positive marginal cost at the origin), which generates an extensive margin (uninformed vs. informed). The parameter c_0 governs *intensity* (the convex marginal cost of improving precision once informed). The linear term $c_1 > 0$ is not strictly necessary to get declining participation. However, testing model predictions about informed and uninformed investors helps to uncover key mechanisms at work. When $c_1 = 0$, all market participants are informed.

Budget constraint. Given participation, information choices, and risky holdings q_j chosen at $t = 1$, terminal wealth is

$$W_{2j} = r\Big(W_{0j} - t_j(F - C(d_j) - p\, q_j)\Big) + q_j z. \tag{3}$$

A nonparticipant does not trade the risky asset, thus $t_j = 0$, $q_j = 0$, $d_j = 0$, and $W_{2j} = rW_{0j}$.

Timing. The economy unfolds over three dates:

$t = 0$ — **Participation and information choice.** Each investor chooses whether to enter and trade (t_j) and, conditional on entry, chooses private-signal precision d_j.

$t = 1$ — **Trading.** Informed participants observe information outcomes s_j. They then submit a menu of price-contingent quantities $q_j(p)$. A competitive market maker posts a price p that clears the market. Participants realized purchases are q_j;

$t = 2$ — **Consumption.** The dividend z is realized and investors consume W_{2j}.

Data needs to be observed before trading for information and its cost to be relevant.

Equilibrium A linear equilibrium consists of (i) pricing rule; (ii) participation and information rules $(t_j, d_j)_{j\in[0,1]}$ chosen at $t = 0$; and (iii) trading rules $(q_j)_{j\in[0,1]}$ at $t = 1$, such that:

1. given the price rule, each investor's choices maximize (1) subject to (3) and the information structure;
2. markets clear for the risky asset for every realization of (z, x);
3. beliefs are Bayesian and consistent with the price rule and the signal structure.

2.2 Solution

Before describing the solution, we define two crucial aggregate variables: (i) *aggregate tolerance* among market participants and (ii) aggregate *tolerance-weighted private data* acquired by the informed sector:

$$\bar{\tau} \;\equiv\; \int t_j \;\tau_j \, dV, \qquad D \;\equiv\; \int \tau_j \, d_j \, dV, \tag{4}$$

where t_j indicates market participation, d_j is investor j's privately chosen signal precision and τ_j is their risk tolerance.

Proposition 1 (Demands and prices)**.** *In any linear rational-expectations equilibrium, the following hold.*

(i) Risky-asset demand. *Conditional on the trading-date information set $\mathcal{I}_j$ (price-only for uninformed, price plus signal for informed), a participating investor's optimal risky demand is*

$$q_j \;=\; \tau_j \, \frac{\mathbb{E}_j[z] - rp}{\mathbb{V}_j[z]}. \tag{5}$$

(ii) Prices and price informativeness. *Market clearing implies a linear price rule*

$$rp \;=\; a \;+\; b_z \, z \;-\; b_x \, x, \tag{6}$$

Appendix A.1.3 reports the coefficients (a, b_z, b_x) and proves that $b_z = D\, b_x$.

The proof of this and all further propositions and results in the text are derived in Appendix A. These are very standard results that arise in many portfolio choice models. We only note that the price coefficient b is what we mean by price informativeness, because it governs the response of prices to changes in fundamental payoffs. This informativeness is increasing in the aggregate information precision D.

The next set of results are the ones that speak to the point of this paper.

Proposition 2 (Information and Participation Decisions)**.** *$C(d_j) = c_1 d_j + \frac{c_0}{2}(d_j)^2$ with $c_0 > 0$. Participation and information acquisition are monotone in wealth and characterized by two thresholds:*

$$W^{\text{part}} = \frac{2\tau F}{\mu^U(D, \bar{\tau})}, \qquad W^{\text{info}} = \frac{2\tau c_1}{\mu(D, \bar{\tau})}. \tag{7}$$

where

$$\mu^U(D, \bar{\tau}) = \frac{\sigma_x^6 \sigma_z^2}{\left(\bar{\tau}\left(D^2\sigma_z^2 + \sigma_x^2\right) + D\,\sigma_x^2\sigma_z^2\right)^2}, \qquad \textit{and} \tag{8}$$

$$\mu(D, \bar{\tau}) = \frac{1 + \mu^U(D, \bar{\tau})}{\sigma_z^{-2} + D^2/\sigma_x^2}. \tag{9}$$

When $c_1 > F$, the population partitions into nonparticipants ($W_{0j} < W^{\text{part}}$), uninformed participants ($W^{\text{part}} \leq W_{0j} < W^{\text{info}}$), and informed participants ($W_{0j} \geq W^{\text{info}}$).

For informed investors, the optimal private precision choice is

$$d_j(W_{0j}) \;=\; \max\left\{0,\; \frac{W_{0j}\mu - 2\tau c_1}{2\tau c_0}\right\}. \tag{10}$$

The functions μ and μ^U are the key objects that explain the interaction between information and participation. To understand them, start with expected utility, at date 0, when information and participation are being chosen.

$$\frac{\tau_j}{2}\,\mathbb{E}\left[S_j^2(\mathcal{I}_j)\right],$$

where $S_j(\mathcal{I}_j)$ is the conditional Sharpe ratio of the risky asset perceived at $t = 1$. So, the expected value of participating in the market is the expected squared Sharpe ratio you believe you will face when investing tomorrow, scaled by risk tolerance, which is increasing in wealth.

The expected squared Sharpe ratio, in turn, is *affine* in information precision d_j, with intercept μ^U and linear coefficient μ:

$$\mathbb{E}\big[\mathcal{S}_j^2(rp, s_j)\big] = \mu^U(D, \bar{\tau}) + \mu(D, \bar{\tau})\, d_j, \tag{11}$$

Thus, μ^U represents the expected squared risk-adjusted return an investor with zero data $d_j = 0$ would face, while μ captures how effective an extra increment of data is in achieving higher risk-adjusted returns.

Since the marginal participant in the market would enter as an uninformed investor, participation depends on the *level* of uninformed opportunities μ^U. In contrast, information demand depends on the *marginal* value μ.

2.3 Financial Technology's Effect on Participation

While technology that reduces entry cost can boost participation, technology that makes information cheaper can reduce participation. When information acquisition scales disproportionately with wealth, the share of informed trading rises, prices become more informative, and the risk-adjusted return available to an uninformed investor falls. This makes entry less appealing, even though average access costs are falling.

Financial technology is a vector of cost shifters (F, c_1, c_0), where F is a fixed market-entry cost, c_1 is the fixed *adoption* cost of becoming informed, and c_0 governs the marginal *intensity* (curvature) cost of information precision. The next proposition explores the effect of each element of the cost vector on participation. In each case, we marginally lower one cost, holding the other two fixed.

Proposition 3 (Technology shocks and participation)**.** *Consider an interior equilibrium in the active-information region* $D > \sigma_x^2/(2\bar{\tau})$.

1. ***Entry-cost reduction (fall in*** F***).*** *The participation cutoff falls:* $W^{\text{part}} \downarrow$.

2. ***Information cost reduction (fall in*** c_1***).*** *The information-adoption cutoff falls (*$W^{\text{info}} \downarrow$*) but the participation cutoff rises (*$W^{\text{part}} \uparrow$*).*

3. ***Reduction in convexity of information cost (fall in*** c_0***).*** *The participation cutoff (*$W^{\text{part}} \uparrow$*) and the information adoption cutoff (*$W^{\text{info}} \uparrow$*) both rise.*

Proof. Part 1 of thes result is mechanical: Entry expands because the fixed cost of market access is lower. Parts 2 and 3 depend on the equilibrium effect on μ^U, the expected squared risk-adjusted return of the uninformed investor.

In part 2, the direct effect of a reduction in information costs c_1 is to reduce the threshold at which informed traders choose to acquire information. We can see this as a reduction in W^{info} in (7). This increase in informed traders, reduces the benefit of being an uninformed trader as μ^U falls in the aggregate amount of information (data) D. From the left side of (7), we see that a decline in μ^U raises the participation threshold W^{part} and drives marginal uninformed agents out of the market.

In part 3, the convexity of information costs c_0 declines. Unlike a reduction in the adoption cost c_1, which draws new investors into the informed sector, a lower c_0 incentivizes those who are already informed to deepen their data acquisition. As shown in the optimal precision choice (10), the precision d_j of informed investors is inversely proportional to c_0. Consequently, as c_0 falls, the most risk-tolerant participants scale up their private information.

This intensification of information acquisition at the top of the wealth distribution raises the aggregate risk-weighted precision, D, which increases asset demand and makes equilibrium prices both higher and more informationally efficient. Both forces reduce the expected squared Sharpe ratio available to an uninformed trader, μ^U, and the

marginal value of adding private data, μ. Because the participation cutoff W^{part} and the information-adoption cutoff W^{info} are both inversely related to these profitability measures, μ_U and μ, both thresholds shift to the right as c_0 falls.

What differs from part 2 is that when the convexity of the information cost falls, fewer investors become informed. However, the effect on participation is similar: Investors exit the market as expected returns fall.

In reality, the effects of lower information costs, of either kind, would show up as high prices, or high price-dividend ratios. Prices rise because well-informed wealthy market participants have high demand. Because assets look expensive and expected returns for the uninformed are low, potential entrants stay out.

2.4 Testable Predictions: Five Hypotheses

The model's mechanism has a simple empirical bite: entry costs and data costs affect participation through different margins. Lower entry costs make it easier to show up in markets; lower data costs change *who becomes informed*, how informative prices are, and therefore the risk-adjusted opportunity set faced by households who trade without information. This delivers five tight hypotheses that guide the empirical design in Section 3.

H1: Participation rises with wealth and with falling market entry costs. This hypothesis, while rather obvious, anchors the analysis in the standard participation-cost view of household finance. Our theory is not contradictory to cost-based explanations. It supports them and augments them with additional information effects.

Both effects are visible in (7), where the participation threshold is a wealth threshold, that is increasing in the entry cost F.

H2: Declining data costs reduce participation in the middle of the wealth distribution. Parts 2 and 3 of Proposition 3 prove that when information costs fall, participation declines because the wealth threshold for participation rises. The rightward shift of the participation threshold W^{part} defined in Equation (7) specifically triggers the exit of the marginal uninformed investors. These are marginal investors who are neither the poorest ($W < W^{part}$), who do not participate, nor the wealthiest ($W > W^{info}$) who benefit from more information.

Cheaper data and analytics do not affect all households symmetrically. The poorest households are typically infra-marginal nonparticipants: fixed costs and low desired risky exposure keep them out regardless. The richest households are infra-marginal participants: they continue to participate and disproportionately scale up information acquisition when data becomes cheaper. The key action is therefore in the middle: households that previously found it worthwhile to participate as price-only investors now face a worse trading environment because informed trading rises and prices become more revealing. As the gains from participating without information compress, the marginal direct participant becomes more likely to exit, concentrating the participation decline in the middle of the distribution.

H3: Declining data costs raise returns for wealthy investors. When data becomes cheaper, wealthy investors increase information acquisition. More information translates into higher realized returns for the wealthy through two channels: better state-contingent portfolio choices (i.e., positions that respond to information) and the ability to capture information rents from trading against less-informed counterparts. At the same time, selection amplifies the pattern: as marginal price-only investors exit, the remaining set of wealthy direct participants is more tilted toward those able to exploit the improved information technology.

This result is implied by the following two results, which break down the mechanisms

by which the wealthy earn those higher returns: more risk-adjusted return and more risk taking. These two channels correspond to the literature on skill vs. risk-taking in financial inequality discussed earlier.

H4: Declining data costs raise risk-adjusted returns (Sharpe ratios) for wealthy investors. As data costs fall, prices incorporate more information, which erodes the risk-adjusted opportunity available to price-only traders. Wealthy investors, in contrast, can convert cheaper data into higher-quality private information and thus preserve—or even improve—their risk-adjusted opportunity set. The model therefore predicts a widening gap in Sharpe ratios: risk-adjusted performance improves for the wealthy (who become more informed) while it compresses for less-informed households, consistent with an adverse-selection interpretation rather than a simple "everyone benefits from lower costs" story.

To show this formally, recognize that there are three groups. Nonparticipants take no risky positions, so $S_N^2 \equiv 0$. Uninformed participants condition only on the equilibrium price and face $\mu^U = \mathbb{E}[S_U^2]$ given by Equation (8). In the active information region, notice that $\partial\mu^U/\partial D < 0$ and $\partial\mu^U/\partial\bar{\tau} < 0$ (or equivalently, $\partial\mu^U/\partial\bar{\tau} < 0$): a more informative price ($D \uparrow$) and a larger risk-bearing sector ($\bar{\tau} \uparrow$) compress the uninformed Sharpe and amplify adverse selection.

Informed participants observe a private signal and choose its precision. Under the quadratic information technology, optimal private precision is defined in Equation (10) above, so the informed Sharpe advantage in Equation (11) becomes

$$\underbrace{\mathbb{E}\big[\mathcal{S}_{I,j}^2(rp, s_j)\big]}_{\text{Informed Sharpe}} = \underbrace{\mu^U(D,\bar{\tau})}_{\text{Uninformed Sharpe}} + \underbrace{\mu(D,\bar{\tau})}_{\text{return to info}} \underbrace{\frac{W_{0j}\mu(D,\bar{\tau}) - 2\tau c_1}{2\tau c_0}}_{\text{Precision}} \tag{12}$$

The expected informed Sharpe advantage is larger than the uninformed squared Sharpe, it is increasing in wealth and becomes more sensitive to wealth when c_0 is smaller. This is the within-informed channel. Notice, also, that the expected informed

Sharpe advantage is decreasing in the aggregate risk-weighted private information precision (D) produced by the informed sector.

Thus, financial technology affects group Sharpe ratios through two channels: (i) it shifts *who* has access to positive Sharpe opportunities (cutoff movements), and (ii) it shifts the *dispersion* of Sharpe opportunities among those who are informed (precision dispersion).

H5: Declining data costs increase wealthy investors' portfolio risk. This prediction may appear contradictory at first glance. After all, the role of better information is to make an asset look less risky, *conditional on the investor's information.* However, from the perspective of an econometrician who does not observe the private signal, this shows up as *greater* unconditional exposure to systematic risk: wealthier, better-informed investors hold larger equity positions and exhibit higher market exposure even if they perceive their own conditional risk to be lower. There are two differences between the investor's risk in an asset and the econometrician's assessment of portfolio risk. One difference is that the asset and the portfolio are not the same. The other difference is that variance is conditional on an information set. The investor and the econometrician have different information.

To reconcile these competing notions of risk formally, consider an informed investor's conditional variance of payoff of the risky asset. A conditional variance is an expected squared prediction error. Investors with more data make smaller prediction errors, on average. By Bayes' law for normal variables, we know that

$$\mathbb{V}(z \mid rp, s_j) = \left(\sigma_{z|p}^{-2} + d_j\right)^{-1} \; < \; \mathbb{V}(z \mid rp) = \sigma_{z|p}^2 \; < \; \mathbb{V}(z) = \sigma_z^2$$

where $\sigma_{z|p}^{-2} \equiv \sigma_z^{-2} + D^2/\sigma_x^2$. This is the sense in which informed investors face less risk. Given what they know, payoffs are more predictable.

The unconditional variance σ_z^2 is the expected squared prediction error of someone

who only knew the normal distribution of z. It also tells us how volatile z is. When an econometrician measures portfolio variance, they aren't measuring an agent's forecast error, they are measuring how volatile the portfolio payoff is. That corresponds to $\mathbb{V}\big(q_j(z-rp)\big)$. It can be high because it is not conditioned on lots of predictive data, because q is high, and because q and $(z-rp)$ covary.

When risk aversion (and thus tolerance τ_j) is positive and the informed investors get data ($d_j > 0$), the informed investors have more variable portfolio payoffs:

$$\mathbb{V}\big(q_j^I(z-rp)\big) - \mathbb{V}\big(q_j^U(z-rp)\big) = \tau_j^2\Big(\mu\, d_j\big(1+4\mu^U\big) + 2(\mu\, d_j)^2\Big) > 0,$$

because μ and μ^u are always positive. Thus, wealthy investors who are also better informed can appear to take on more systematic risk in unconditional data, even though their *conditional* risk per share is lower.

Corollary 1 (A one-factor "effective beta")**.** *In this one-risky-asset economy, a natural one-factor measure of systematic exposure is the variance ratio*

$$\beta_j^{\text{eff}} \;\equiv\; \sqrt{\frac{\mathbb{V}\big(q_j(z-rp)\big)}{\mathbb{V}(z-rp)}}\,,$$

where $\mathbb{V}(z-rp)$ is the unconditional variance of the market excess payoff (observable in returns data). Since $\mathbb{V}(z-rp)$ is common across investors, the ordering of β_j^{eff} mirrors the ordering of $\mathbb{V}(q_j(z-rp))$: informed investors (with $d_j > 0$) have strictly larger β_j^{eff} than uninformed investors. Higher risk tolerance further increases β_j^{eff}.

Note that wealthier households choose higher risk exposure β_j^{eff} simply because they are more risk tolerant. However, even holding wealth and risk tolerance constant, more informed investors also choose higher β_j^{eff}. This more subtle prediction is what we later test.

2.5 What About Funds?

What if investors can pool resources and share the cost of information acquisition? The participation paradox survives when households can *delegate* information production and benefit from pooling. For two reasons, one empirical and one theoretical.

Empirically, pooling is not frictionless. In practice, access to delegated vehicles with lower fees or more attractive terms is not uniform across households, but is often tiered by account size. Even within the same fund, institutional share classes commonly require much larger minimum investments and charge lower fees than retail classes. For example, in the T. Rowe Price Large-Cap Growth Fund, the Investor Class generally requires a $2,500 minimum initial investment and carries a 0.70% net expense ratio after fee waivers, whereas the corresponding I Class requires a $500,000 minimum initial investment and carries a 0.55% net expense ratio. Thus, the lower-cost version of the same product is effectively reserved for larger investors. This suggests that delegation does not necessarily eliminate the participation paradox: although pooling can spread some information costs, access to the most attractive delegated products remains easier for larger investors.

Theoretically, the puzzle survives because investors have to search for funds (Gârleanu and Pedersen, 2018). Suppose investors who enter to either (i) trade uninformed, (ii) acquire private precision individually, or (iii) pay a search/participation cost F_m to join a competitively supplied fund that produces a *common* signal for its investors. Pooling makes information partly a club good: by spreading fixed components of information costs across many investors, delegation can turn some marginal *uninformed* participants into *informed* participants. At the top, very risk-tolerant investors still prefer to self-produce and tailor information intensity to their own position size, generating a natural "delegation window."

Importantly, delegation changes *how* information is purchased, not the equilibrium force that links information to participation. Delegated learning raises aggregate pri-

vate precision and therefore price informativeness, which compresses the risk-adjusted return available to uninformed traders through the same adverse-selection channel emphasized in the baseline. As a result, improvements that lower information costs (or lower the cost of accessing pooled information) can still reduce direct participation at the margin by eroding the gains from trade for less-informed households. Appendix A.4 presents the delegated-learning environment, derives the associated cutoffs (participation, delegation, and self-learning), and quantifies how the main comparative statics mirror the benchmark results.

2.6 What About Dynamics?

Adding dynamic elements to the model does not overturn the core adverse-selection mechanism because the primary driver remains the relative gain for marginal households entering as less-informed participants. In a dynamic model, this force is strongest precisely where the data suggest it should be: among middle-wealth households that are close to the participation margin, not among households that are always out of the market or among wealthy households that remain infra-marginal participants. Appendix A.5 presents this result. Standard household finance frictions—such as participation persistence, reentry costs, labor income risk, and experience-based beliefs—serve to amplify this equilibrium force rather than replace it. For instance, persistence makes the effects of a deteriorating uninformed opportunity set more long-lived because staying in the market is cheaper than reentering after an exit. Ultimately, these dynamic factors ensure that when information technology favors those with scale, direct market participation becomes less attractive for those without it, confirming that the model's central mechanism remains the same.

3 Empirical Evidence of Testable Implications

This section evaluates the model's main cross-sectional predictions. The mechanism developed above implies that financial technology affects both who participates in risky asset markets and, conditional on participation, which investors become informed. On the participation margin, the model predicts that participation rises with wealth and that declines in data costs reduce participation particularly in the middle of the wealth distribution. Among those who do participate, investors with larger risky positions have stronger incentives to become informed. As a result, they should earn higher returns, achieve higher risk-adjusted returns, and hold portfolios that appear riskier in unconditional data. We test these predictions using two complementary datasets. Using SCF data, we study participation patterns across the wealth distribution, speaking to Hypotheses H1 and H2. Using portfolio-level data from Addepar, which allows us to compare returns across investors that differ in wealth and in the sophistication of the intermediaries they use, we test Hypotheses H3 through H5.

3.1 Data and empirical design

Our empirical analysis complements the SCF with portfolio-level data from Addepar, a portfolio accounting and reporting platform used by a broad set of wealth managers and advisors, including registered investment advisors, brokers, family offices, and private banks. The platform has reached a broad coverage of households' financial activities covering over $8 trillion in assets and more than 1,300 financial advisors. The data provide monthly holdings and returns at the security level across a wide set of asset classes, including both publicly traded securities and illiquid private asset classes, between 2016 and 2024. This breadth makes the data particularly well suited to studying heterogeneity in household portfolio choice and investment outcomes.

For the purposes of this paper, we focus on the subset of liquid asset classes, used

in Gabaix, Koijen, Mainardi, Oh and Yogo (2024) and Mainardi (2025). In practice, this includes publicly traded securities. We make this restriction because private asset classes are subject to additional frictions such as illiquidity, infrequent trading, stale valuations, and institutional access constraints that are not modeled directly in our framework. These frictions are economically important in private-asset settings as shown in Balloch et al. (2025). Focusing on liquid assets therefore allows for a cleaner empirical mapping between the model and the data.

At the investor-month level, we observe security-level holdings and returns. We construct returns on liquid assets as the value-weighted average return across all liquid securities held by investor i in month t. We also construct returns on risky assets, which exclude U.S. treasuries and municipal bonds, which we treat as the empirical counterpart of the risk-free asset in the model. This distinction matters because the model's participation and portfolio-allocation decisions operate over the risky share of the portfolio.

A key empirical challenge is to identify which investors are more likely to correspond to the informed investors in the model. Guided by prior work, we classify as informed those investors advised by single family offices and private banks. The reason is that both types of intermediaries have been shown to behave more closely to rational or informed benchmarks in different settings. In particular, Gabaix et al. (2024) show that while the majority of investors in Addepar trade procyclically, households advised by single family offices trade countercyclically, as prescribed by standard rational benchmarks. Moreover, Mainardi (2025) shows that investors advised by private banks trade more closely to rational benchmarks in the presence of frictions, in particular capital gains taxes. We therefore use intermediary sophistication as our main proxy for informed trading in the Addepar data.

3.2 Participation across the wealth distribution (H1, H2)

Hypotheses H1 and H2 were that wealthier households participate more and that the decline in participation was concentrated in the middle-income households. Figure 2 repeats the analysis from Section 1, but now reports participation rates separately by wealth quintile. For simplicity, we focus on participation in risky assets outside retirement accounts, which maps more directly to the model because it reflects households' active decisions to enter risky asset markets.

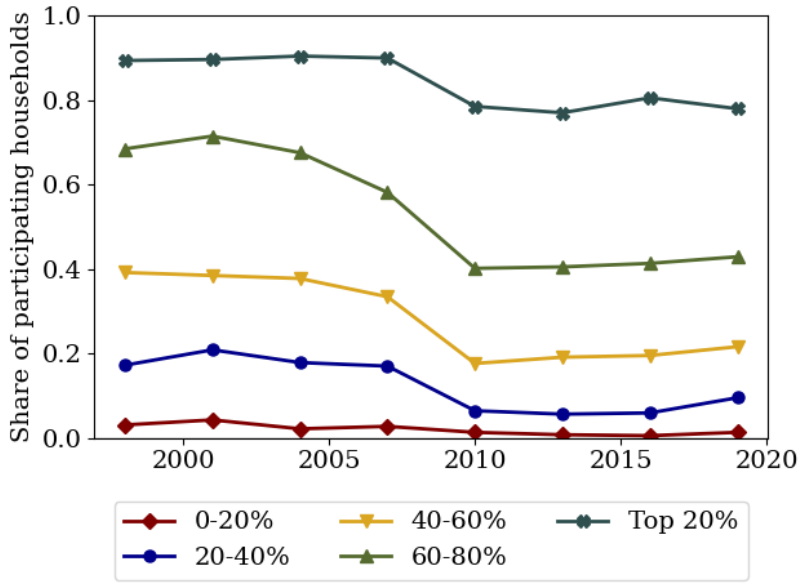


Figure 2: Participation in risky assets outside retirement accounts by wealth quintile. This figure plots participation rates in risky assets outside retirement accounts separately by wealth quintile using SCF data.

We find evidence consistent with the model's predictions. First, participation rises strongly with wealth: wealthier households are disproportionately more likely to hold risky assets outside retirement accounts, consistent with H1. Second, as data costs continued to decline throughout the 2000s, the fall in participation was concentrated in the three middle quintiles of the wealth distribution, where participation rates fell by roughly one half. By contrast, changes are more moderate at the very bottom and the very top of the wealth distribution. This pattern is consistent with H2 and closely

matches the model's mechanism: households at the bottom are largely infra-marginal nonparticipants, households at the top are infra-marginal participants, and the main adjustment occurs in the middle of the distribution, where households are most sensitive to the deterioration in the uninformed gains from market participation.

While these predictions are consistent with any theory of declining benefits of rising costs, they set the stage for results more specific to our mechanism.

3.3 Returns across investor types (H3)

Hypothesis H3 predicts that wealthier and more informed investors earn higher returns. Using the Addepar data, we classify investors into five quintiles based on total portfolio wealth. We then report in Table 1 average monthly returns on risky assets across investors and year-months for each wealth quintile, as well as the difference in average returns between informed and uninformed investors within each quintile.

Table 1: Average returns by wealth quintile and informed status on risky assets

	0–20%	20–40%	40–60%	60–80%	Top 20%
Average returns	0.0079	0.0078	0.0080	0.0084	0.0087
Informed	0.0004	0.0010	0.0005	0.0004	0.0001

Notes: This table reports average monthly returns by wealth quintile. The row *Informed* reports the difference in average monthly returns between informed and uninformed investors within each wealth quintile, where informed investors are defined as those advised by single family offices or private banks.

Average returns increase with wealth, rising from 79 basis points per month in the bottom quintile to 87 basis points in the top quintile. This 8-basis-point gap at monthly frequency corresponds to 96 basis points at annual frequency and is therefore economically meaningful. Moreover, within each wealth quintile, informed investors earn higher returns than uninformed investors.

3.4 Risk-adjusted returns across investor types (H4)

We then test Hypothesis H4 that wealthier and more informed investors earn higher risk-adjusted returns. To this end, within each quintile, we estimate a panel factor model of monthly portfolio returns. Specifically, for each wealth quintile, we regress investor-level returns on the standard five Fama-French factors (the market factor, HML, SMB, CMA, RMW) plus momentum. We include in the regression an indicator for informed investors and we cluster standard errors at the year-month level to account for cross-sectional residual correlation.

Table 2 reports the results when using returns on risky assets on the left hand-side of the factor model. Consistent with the model, alpha increases monotonically with wealth, rising from -27 basis points in the bottom quintile to -21 basis points in the top quintile at monthly frequency, corresponding to a gap of 72 basis points at annual frequency. In other words, investors with greater scale earn systematically higher risk-adjusted returns. This is exactly what the model predicts: because information is more valuable when it can be applied to larger risky positions, wealthier investors have stronger incentives to acquire precision and therefore earn higher returns after accounting for risk.

The table also shows that, within a given wealth quintile, informed investors tend to outperform uninformed investors by up to 14 basis points at monthly frequency, equivalent to 1.7 percentage points at annual frequency. This effect is strongest in the lower three wealth quintiles and becomes smaller toward the very top of the wealth distribution. That pattern is also natural in the model. At the top of the wealth distribution, a larger fraction of investors are likely above the informed wealth threshold, so differences across intermediary types are smaller. By contrast, in the middle and lower parts of the wealth distribution, access to more sophisticated intermediation is more likely to separate investors who are able to exploit information from those who are not.

Table 2: Alpha by wealth quintile and informed status on risky assets

	0–20%	20–40%	40–60%	60–80%	Top 20%
α	-0.0027***	-0.0025***	-0.0024***	-0.0023***	-0.0021***
	(0.0006)	(0.0006)	(0.0006)	(0.0005)	(0.0005)
Informed	0.0011**	0.0014**	0.0010*	0.0009**	0.0002
	(0.0005)	(0.0006)	(0.0005)	(0.0004)	(0.0002)
R^2	0.6237	0.6754	0.7027	0.7109	0.6595
Observations	1,627,817	1,629,467	1,632,680	1,629,411	1,627,962

Notes: This table reports estimates from monthly panel regressions of investor excess returns on risky assets on an indicator for informed investors, estimated separately by wealth quintile. Risky assets exclude U.S. treasuries and municipal bonds. All regressions also include the CRSP market factor and the standard Fama–French five factors plus momentum (HML, SMB, CMA, RMW, and momentum), whose coefficients are omitted for brevity. The variable *Informed* is an indicator for investors advised by single family offices or private banks. Standard errors, clustered by year-month, are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The negative alphas should not be surprising. These are households, not professionals. Only professionals earn positive alpha, sometimes. When they outperform, someone else must underperform for the market to clear. What is important is that richer and more informed households have alphas that are less negative. They earn more.

Because the model predicts that wealthier and informed investors outperform also on their overall portfolio, Table 3 repeats the analysis using returns on all liquid assets. We find similar results: the gap between the bottom and top quintile is 5 basis points at monthly frequency, equivalent to 60 basis points at annual frequency. Moreover, informed investors in the lower wealth quintiles outperform uninformed investors by up to 1.2 percentage points at annual frequency.

These alpha differentials are economically significant. Focusing on returns on financial wealth, Fagereng et al. (2020) estimate a gap of 1.5 percentage points between the top and bottom of the Norwegian wealth distribution, with about 60 basis points attributable to differences in returns on risky assets. Bach, Calvet and Sodini (2020) report a 55-basis-point alpha gap in Sweden, again for financial wealth, similar to the

Table 3: Alpha by wealth quintile and informed status on all liquid assets

	0–20%	20–40%	40–60%	60–80%	Top 20%
α	-0.0027***	-0.0025***	-0.0024***	-0.0023***	-0.0022***
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Informed	0.0010**	0.0010*	0.0006	0.0005	0.0001
	(0.0005)	(0.0006)	(0.0005)	(0.0004)	(0.0002)
R^2	0.6254	0.6637	0.6826	0.6757	0.6062
Observations	1,647,727	1,658,569	1,662,121	1,664,034	1,664,297

Notes: This table reports estimates from monthly panel regressions of investor excess returns on all liquid assets on an indicator for informed investors, estimated separately by wealth quintile. All regressions also include the CRSP market factor and the standard Fama–French five factors plus momentum (HML, SMB, CMA, RMW, and momentum), whose coefficients are omitted for brevity. The variable *Informed* is an indicator for investors advised by single family offices or private banks. Standard errors, clustered by year-month, are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

57-basis-point estimate in Campbell, Ramadorai and Ranish (2019). The alpha differentials documented here are directly comparable to and sometimes larger than these earlier estimates.

3.5 Risk taking across investor types (H5)

Our fifth hypothesis was that informed investors' portfolios should load more on market risk. To test this, we extend the factor-model above: We estimate the return regression separately for each wealth quintile × informed-status group so that we can report differences in the factor loadings by group.

Figure 3 plots the estimated coefficients on the market factor separately for informed and uninformed investors within each wealth quintile. Within each wealth group, informed investors have higher market beta than uninformed investors. Thus, the same investors who earn higher risk-adjusted returns also hold portfolios with greater systematic risk exposure.

This finding is tightly connected to the model. In the theory, more informed investors do not take on more unconditional risk because they are mechanically more

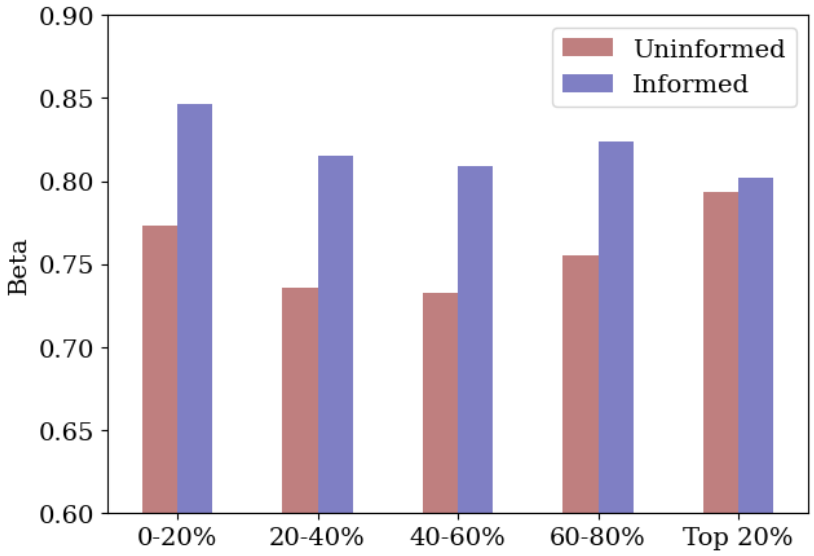


Figure 3: Market beta by wealth quintile and informed status. This figure plots the estimated loading on the market factor from monthly return regressions estimated separately for each wealth-quintile × informed-status. Within each wealth group, informed investors exhibit higher market beta than uninformed investors. This pattern is consistent with the model: investors with superior information optimally allocate more to risky securities, so to an econometrician who does not observe their information set, they appear to take on more systematic risk.

risk-loving. That would explain higher risk taking by wealthier investors, but not more risk taking by informed investors in the same wealth category. They take on more risk because, conditional on their superior information, risky assets are more predictable and therefore less uncertain and less risky, to them. An econometrician who does not observe the investor's information set therefore sees higher unconditional factor exposure and interprets it as greater risk-taking.

This evidence also helps reconcile two influential views in the household-finance literature regarding why wealthier households earn higher returns. Bach et al. (2020) emphasize that wealthier households earn higher returns because they take on more systematic risk. Fagereng et al. (2020) emphasize instead persistent skill differences across households. Our model suggests that these mechanisms need not be in contradiction. Wealthier households are more informed and therefore more skilled, but their payoffs are also more variable. In that sense, the skill channel and the risk channel are

two manifestations of the same underlying force: differential access to information.

4 Quantification: Does the Evidence Explain Trends in Participation?

We provide a compact welfare illustration designed to quantify the mechanism, not to estimate it. The goal is to show that information technology can generate sizable participation responses and meaningful welfare shifts across the distribution through Sharpe compression and selection out of the market.

Table 4: Information technology, participation, Sharpe compression, and welfare (lognormal illustration)

	Baseline	F reduced	c_1 reduced	c_0 reduced
Participation rate $\Pr(W_0 \geq W^{\text{part}})$	58.9%	88.9%	53.2%	32.8%
Uniformed share $\Pr(W^{\text{part}} \leq W_0 \leq W^{\text{info}})$	40.3%	70.3%	23.8%	25.5%
Informed share $\Pr(W_0 \geq W^{\text{info}})$	18.6%	18.6%	29.4%	7.3%
Participation wealth threshold W^{part}	1443.8	532.9	1667.4	2826.0
Informed wealth threshold W^{info}	4414.9	4414.9	3105.5	7740.1
Uninformed squared Sharpe μ^U	$1.7e^{-17}$	$1.4e^{-17}$	$1.4e^{-17}$	$8.5e^{-18}$
Informed squared Sharpe $\mathbb{E}\big[\mathcal{S}^2_{I,j}(rp, s_j)\big]$	$1.7e^{-06}$	$1.7e^{-06}$	$1.2e^{-06}$	$2.4e^{-06}$
Marginal value of data μ	$3.8e^{-05}$	$3.8e^{-05}$	$3.4e^{-05}$	$2.2e^{-05}$
Agg. risk-bearing capacity $\bar{\tau}$	2946.9	3268.6	2848.9	2351.6
Agg. risk-weighted private data D	162.0	162.0	170.8	214.5

Notes: Initial wealth and primitives are held fixed across scenarios (maximum wealth is 186,249, median wealth is 1,812, mean wealth is 2,998). In each case we solve a fixed point for $(D, \bar{\tau})$. The baseline scenario targets a 58.9% participation share, and a 18.6% informed share and a mean-preserving normalization of average terminal wealth to keep the numbers comparable across scenarios. Each cost reduction is at 70% from the initial level with the exception of c_1 for which we explicitly target the increase in the informed share in the SCF between 1998 and 2019.

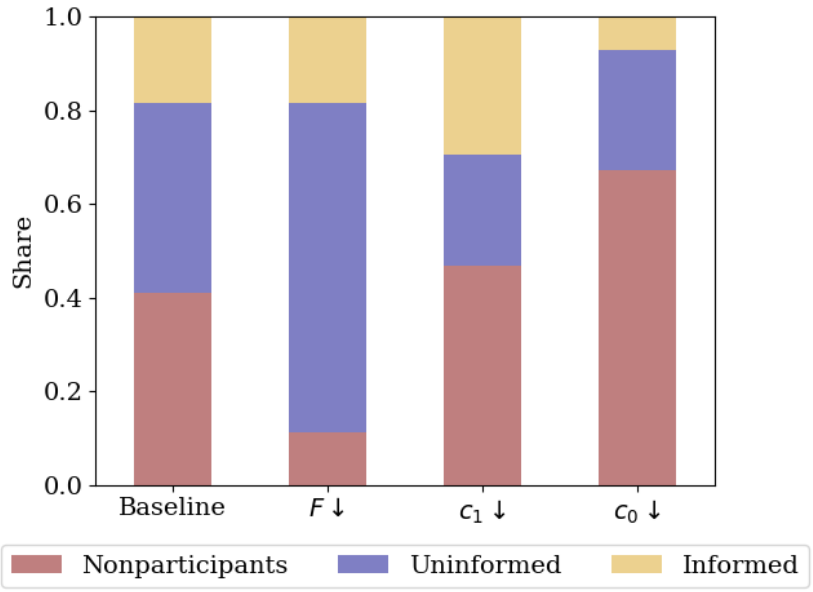
Design. We assume a lognormal initial wealth distribution $W_0 \sim \log \mathcal{N}(\mu, \sigma^2)$ and choose (F, c_1, c_0) to match baseline moments for participation and the informed share (reported in Table 4) as of 1998 from the SCF. We define the share of informed investors as the fraction of households that report using a professional financial planner or advisor.

We then implement one-at-a-time same-magnitude reductions in participation costs ($F \downarrow$) , information adoption costs ($c_1 \downarrow$) and information intensity costs ($c_0 \downarrow$), holding fixed all other primitives. We are particularly interested in evaluating the effect of decreasing c_1 on the participation rate so we change c_1 to target the increase in the informed share from 18.6% to 29.4% between 1998 and 2019 in the SCF.

Results. In the calibrated economy, lowering information costs raises price informativeness and reduces the uninformed Sharpe ratio μ_U. Participation declines because marginal households no longer find it worthwhile to pay the fixed entry cost for a compressed risk-adjusted return. Welfare effects are correspondingly heterogeneous: households that remain uninformed (or exit) lose surplus, while wealthy informed investors capture the gains from cheaper precision and information rents. Interestingly, while the model can match the increase in the informed share by design, the simulated decrease in c_1 implies a decrease in the participation rate from 58.9% to 53.2%. This is well-aligned with the decrease from 58.9% to 53.1% observed between 1998 and 2019 in the SCF. This exercise therefore suggests that our mechanism is powerful and can account for a large fraction of the decrease in the participation rate observed over the last two decades.

Figure 4 displays this mechanism visually. It isolates the extensive-margin reallocations across groups (nonparticipants, uninformed participants, informed). A reduction in the participation cost F primarily shifts mass from nonparticipation into uninformed participation, while a reduction in the adoption cost c_1 shifts mass from uninformed to informed, but also from the uninformed to nonparticipants through the adverse selection channel emphasized above.

Figure 4: **Quantitative decomposition: extensive vs. intensive margins.** Entry technology expands participation. Information technology reduces participation.



Notes. The figure reports the endogenous mass shares of nonparticipants, uninformed participants, and informed participants across the baseline (initial separation of nonparticipants, uninformed, and informed participants) and one-at-a-time cost cuts in (F, c_1, c_0), respectively, under a lognormal initial wealth distribution.

5 Conclusion

The participation paradox reveals that financial democratization stalled because technological progress is not a neutral force for market inclusion. By distinguishing between trading technologies that lower entry barriers and data provision that sharpens informational advantages, our novel three-layered model demonstrates that falling data costs trigger an equilibrium adverse-selection channel. Because the benefits of information scale with portfolio size, wealthier investors utilize cheaper analytics to disproportionately increase their private precision, which makes prices more informationally efficient but simultaneously compresses the risk-adjusted returns available to uninformed households. This mechanism explains why the post-2005 retrenchment in participation was concentrated among middle-wealth households, a fact we verify using SCF data. Furthermore, our measurement using granular Addepar portfolio data confirms that both

alpha and systematic risk exposure rise with wealth and intermediary sophistication, providing a disciplined reconciliation of long-standing debates over skill versus risk-taking in financial inequality.

These insights suggest that the distributional consequences of modern financial technology hinge on who can effectively transform abundant data into actionable information at scale. A central conclusion of our quantitative exercise is that lowering entry costs alone is insufficient to offset the deterrent effect of adverse selection created by the data revolution. Consequently, an important next step for research and policy is to explore interventions that encourage participation by guarding smaller investors against being picked off by an informed elite. Such measures might include improving access to sophisticated delegated vehicles to the point where the marginal participant is as well-informed as institutional actors. By explicitly addressing the tension between informational efficiency and participation breadth, policymakers can better ensure that technological innovation broadens the risk-sharing capacity of the economy.

T

References

Azarmsa, Ehsan, "Investment Sophistication and Wealth Inequality," SSRN Working Paper No. 3206823 2022.

Bach, Laurent, Laurent E Calvet, and Paolo Sodini, "Rich pickings? Risk, return, and skill in household wealth," *American Economic Review*, 2020, *110* (9), 2703–2747.

Balloch, Cynthia, Federico Mainardi, Sangmin Oh, and Petra Vokata, "Democratizing Private Markets: Private Equity Performance of Individual Investors," Working Paper 5319498, Columbia Business School June 2025.

Benhabib, Jess, Alberto Bisin, and Mi Luo, "Earnings Inequality and Other Determinants of Wealth Inequality," *American Economic Review*, May 2017, *107* (5), 593–97.

Bretscher, Lorenzo, Francisco J. Gomes, Riccardo Sabbatucci, and Andrea Tamoni, "Return Heterogeneity in Retirement Accounts," Technical Report, SSRN, Stockholm, Sweden 2022.

Calvet, Laurent E, Laurent Bach, and Paolo Sodini, "Rich Pickings? Risk, Return, and Skill in the Portfolios of the Wealthy," HEC Research Papers Series 1126, HEC Paris January 2019.

Campbell, John Y, Tarun Ramadorai, and Benjamin Ranish, "Do the rich get richer in the stock market? Evidence from India," *American Economic Review: Insights*, 2019, *1* (2), 225–240.

Cuoco, Domenico and Ron Kaniel, "Equilibrium Prices in the Presence of Delegated Portfolio Management," *Journal of Financial Economics*, 2011, *101* (2), 264–296.

Dessaint, Olivier, Thierry Foucault, and Laurent Frésard, "Does Alternative Data Improve Financial Forecasting? The Horizon Effect," *Journal of Finance*, 2024, *79* (3), 2237–2287.

Easley, David, Maureen O'Hara, and Liyan Yang, "Differential Access to Price Information in Financial Markets," *Journal of Finance*, 2016, *71* (3), 1071–1112.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, "Heterogeneity and Persistence in Returns to Wealth," *Econometrica*, 2020, *88* (1), 115–170.

Favilukis, Jack, "Inequality, Stock Market Participation, and the Equity Premium," *Journal of Financial Economics*, 2013, *107* (3), 740–759.

Frost, Jon, Leonardo Gambacorta, and Romina Gambacorta, "On the nexus between wealth inequality, financial development and financial technology," *Journal of Economic Behavior & Organization*, 2022, *202*, 429–451.

Gabaix, Xavier, Ralph SJ Koijen, Federico Mainardi, Sangmin Oh, and Motohiro Yogo, "Asset demand of us households," Technical Report, National Bureau of Economic Research 2024.

Gambacorta, Leonardo, Romina Gambacorta, and Roxana Mihet, "FinTech, Investor Sophistication, and Financial Portfolio Choices," *The Review of Corporate Finance Studies*, 06 2023, *12* (4), 834–866.

Gârleanu, Nicolae and Lasse H. Pedersen, "Efficiently Inefficient Markets for Assets and Asset Management," *The Journal of Finance*, 2018, *73* (4).

Gomes, Francisco J. and Alexander Michaelides, "Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence," *The Journal of Finance*, April 2005, *60* (2), 869–904.

_ , Michael Haliassos, and Tarun Ramadorai, "Household Finance," *Journal of Economic Literature*, 2021, *59* (3), 919–1000.

He, Zhiguo and Wei Xiong, "Delegated Asset Management, Investment Mandates, and Asset Prices," *Journal of Monetary Economics*, 2013, *60* (5), 571–590.

Kacperczyk, Marcin, Jaromir Nosal, and Luminita Stevens, "Investor sophistication and capital income inequality," *Journal of Monetary Economics*, 2018.

Katona, Zsolt, Marcus Painter, Panos Patatoukas, and Jean Zeng, "On the Capital Market Consequences of Big Data: Evidence from Outer Space," *Journal of Financial and Quantitative Analysis*, 2025, *60* (2), 551–579.

Mainardi, Federico, "Who Harvests? Tax Alpha and Heterogeneous Responses to Capital Gains Taxation," *Tax Alpha and Heterogeneous Responses to Capital Gains Taxation (October 31, 2025)*, 2025.

Peress, Joël, "Wealth, Information Acquisition, and Portfolio Choice," *Review of Financial Studies*, 2004, *17* (3), 879–914.

_ , "Information vs. Entry Costs: What Explains U.S. Stock Market Evolution?," *Journal of Financial and Quantitative Analysis*, September 2005, *40* (03), 563–594.

Poterba, James M., "The Changing Pattern of Stock Ownership in the U.S.: 1989–2013," *Eastern Economic Journal*, 2018, *44* (1), 1–17.

Smith, Matthew, Owen Zidar, and Eric Zwick, "Top Wealth in America: New Estimates and Implications for Taxing the Rich," Working Paper 29374, NBER 2021.

Vissing-Jørgensen, Annette, "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution," *Journal of Political Economy*, August 2002, *110* (4), 825–853.

Vissing-Jorgensen, Annette, "Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures," 2002. FED Working Paper 8884.

Wolff, Edward N., "Household Wealth Trends in the United States, 1962-2013: What Happened over the Great Recession?," Working Paper 20733, NBER December 2014.

_ , *A Century of Wealth in America*, Cambridge, MA and London, England: Harvard University Press, 2017.

A Theoretical model

A.1 Equilibrium derivations, Sharpe ratios, and wealth cutoffs

Primitives and timing. There is a riskless asset with gross return $r > 0$ and one risky asset with terminal payoff $z \sim \mathcal{N}(\mu_z, \sigma_z^2)$. Noise-trader supply is $x \sim \mathcal{N}(\mu_x, \sigma_x^2)$, independent of z. At date 0, household j with initial wealth W_{0j} decides whether to enter the risky-asset market by paying an entry fee $F \geq 0$. Conditional on entry, the household can acquire a private signal

$$s_j \;=\; z + \varepsilon_j, \qquad \varepsilon_j \sim \mathcal{N}(0, d_j^{-1}), \qquad \varepsilon_j \perp (z, x),$$

at cost

$$C(d_j) \;=\; c_1\, d_j \;+\; \frac{c_0}{2}\big(d_j\big)^2, \qquad c_1, c_0 > 0.$$

At date 1, the price is observed; informed investors additionally observe s_j. They choose their position q_j in the risky asset. At date 2, z realizes.

Preferences. Households have mean–variance preferences with wealth-dependent absolute risk tolerance

$$\tau(W_{0j}) \;\equiv\; \frac{W_{0j}}{\tau}, \qquad \tau > 0.$$

Let rp denote the risky asset's *gross* price at date 1.

A.1.1 Portfolio demand

If household j participates and chooses precision d_j, then terminal wealth is

$$W_{2j} = r\Big(W_{0j} - F - C(d_j)\Big) + q_j\,(z - rp).$$

Given an information set $\mathcal{I}_j$ at date 1, household j solves

$$\max_{q_j} \Big\{\mathbb{E}[z - rp \mid \mathcal{I}_j]\, q_j - \frac{\tau}{2W_{0j}}\mathbb{V}(z - rp \mid \mathcal{I}_j)\, q_j^2\Big\}.$$

The first-order condition yields the usual mean–variance demand:

$$q_j(\mathcal{I}_j) = \tau_j\, \frac{\mathbb{E}[z \mid \mathcal{I}_j] - rp}{\mathbb{V}(z \mid \mathcal{I}_j)}. \tag{13}$$

Plugging the optimizer back in gives the indirect (date-1) value of trading:

$$\max_{q_j}\Big\{\mathbb{E}[z - rp \mid \mathcal{I}_j]\, q_j - \frac{\tau}{2W_{0j}}\mathbb{V}(z - rp \mid \mathcal{I}_j)\, q_j^2\Big\} = \frac{W_{0j}}{2\tau}\,\mathcal{S}^2(\mathcal{I}_j), \tag{14}$$

where the (date-1) *squared Sharpe ratio* is

$$\mathcal{S}^2(\mathcal{I}_j) \;\equiv\; \frac{\big(\mathbb{E}[z \mid \mathcal{I}_j] - rp\big)^2}{\mathbb{V}(z \mid \mathcal{I}_j)}. \tag{15}$$

A.1.2 Linear price, posteriors, and price informativeness

We focus on a rational expectations equilibrium with a linear price rule

$$rp \;=\; a + b_z z - b_x x, \qquad (a, b_z, b_x) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+. \tag{16}$$

Posterior given price only. Because (z, x, rp) is jointly normal under (16), the posterior of z given rp is normal with

$$\mathbb{E}[z \mid rp] = \mu_z + \kappa\,(rp - \mathbb{E}[rp]), \qquad \kappa \equiv \frac{(z, rp)}{\mathbb{V}(rp)} = \frac{b_z \sigma_z^2}{b_z^2 \sigma_z^2 + b_x^2 \sigma_x^2}, \tag{17}$$

$$\mathbb{V}(z \mid rp) = \sigma_z^2 - \frac{(z, rp)^2}{\mathbb{V}(rp)} = \frac{1}{\sigma_z^{-2} + \left(\frac{b_z}{b_x}\right)^2 \sigma_x^{-2}}. \tag{18}$$

Define the *posterior precision after observing price*

$$\sigma_{z|p}^{-2} \;\equiv\; \frac{1}{\mathbb{V}(z \mid rp)} = \sigma_z^{-2} + \left(\frac{b_z}{b_x}\right)^2 \sigma_x^{-2}. \tag{19}$$

Posterior given price and a private signal. If investor j also observes $s_j = z + \varepsilon_j$ with precision d_j, then

$$\mathbb{E}[z \mid rp, s_j] = \mathbb{E}[z \mid rp] + \omega_j\big(s_j - \mathbb{E}[z \mid rp]\big), \qquad \omega_j \equiv \frac{d_j}{\sigma_{z|p}^{-2} + d_j}, \tag{20}$$

$$\mathbb{V}(z \mid rp, s_j) = \frac{1}{\sigma_{z|p}^{-2} + d_j}. \tag{21}$$

A.1.3 Aggregates and equilibrium price coefficients

Let $\tau(W_0) = W_0/\tau$. Let V denote the (endogenous) cross-sectional distribution of households among participants (after entry and information choices). Define the two equilibrium aggregates:

$$\bar{\tau} \equiv \int t_j \tau(W_0)\, dV, \qquad D \equiv \int \tau(W_0)\, d_j\, dV, \tag{22}$$

where t_j is the trade decision (i.e, 0 for non-participants), $d_j = 0$ for uninformed participants (so the integral for D effectively runs over informed households).

Aggregate demand. Using (13) and (20)–(21), each informed investor's demand can be written as

$$q_j = \tau(W_{0j})\Big(\sigma_{z|p}^{-2} + d_j\Big)\big(\mathbb{E}[z \mid rp] - rp\big) + \tau(W_{0j})\, d_j\big(s_j - \mathbb{E}[z \mid rp]\big),$$

while an uninformed investor has the same expression with $d_j = 0$. Aggregating across a continuum of investors, the idiosyncratic signal noises average out, so

$$\int q_j \, dV = \Big(\sigma_{z|p}^{-2}\,\bar{\tau} + D\Big)\big(\mathbb{E}[z \mid rp] - rp\big) + D\,\big(z - \mathbb{E}[z \mid rp]\big). \tag{23}$$

Market clearing requires $\int q_j \, dV = x$.

Proposition 4 (Equilibrium price coefficients)**.** *In a linear equilibrium* (16) *with aggregates* $(D, \bar{\tau})$*, the slope coefficients satisfy*

$$b_z \;=\; D\, b_x, \qquad b_x \;=\; \frac{\sigma_z^2\big(D\,\bar{\tau} + \sigma_x^2\big)}{\bar{\tau}\big(D^2\sigma_z^2 + \sigma_x^2\big) + D\,\sigma_x^2\sigma_z^2}. \tag{24}$$

The intercept a is pinned down by expected market clearing:

$$a = \mu_z(1 - b_z) + b_x\mu_x - \frac{\mu_x}{\sigma_{z|p}^{-2}\,\bar{\tau} + D}, \qquad \sigma_{z|p}^{-2} = \sigma_z^{-2} + \frac{D^2}{\sigma_x^2}. \tag{25}$$

Proof. Insert (17) into the market-clearing condition $\int q_j \, dV = x$ using (23). Write $rp - \mathbb{E}[rp] = b_z(z - \mu_z) - b_x(x - \mu_x)$ from (16). Because x and z are independent and the equilibrium is linear-normal, market clearing must hold state-by-state; matching coefficients on $(z - \mu_z)$ and $(x - \mu_x)$ yields the unique solution (24). Taking expectations of (23) and using $\mathbb{E}[\mathbb{E}[z \mid rp]] = \mu_z$ gives $\mu_x = (\sigma_{z|p}^{-2}\bar{\tau} + D)\big(\mu_z - \mathbb{E}[rp]\big)$, which implies (25) after substituting $\mathbb{E}[rp] = a + b_z\mu_z - b_x\mu_x$. Finally, (19) together with $b_z/b_x = D$ delivers $\sigma_{z|p}^{-2} = \sigma_z^{-2} + D^2/\sigma_x^2$. □

Interpretation. $\bar{\tau}$ summarizes total risk-bearing capacity of market participants; D summarizes the *risk-tolerance-weighted* amount of private information. Price informativeness about z is governed by $\sigma_{z|p}^{-2} = \sigma_z^{-2} + D^2/\sigma_x^2$: more private information in the economy (larger D) makes the price more informative.

A.1.4 Sharpe ratios for uninformed and informed investors

Uninformed (price-only) realized Sharpe. An uninformed participant has information set $\mathcal{I}_U = \sigma(rp)$ at date 1, hence

$$\mathcal{S}_U^2(rp) = \frac{\big(\mathbb{E}[z \mid rp] - rp\big)^2}{\mathbb{V}(z \mid rp)} = \sigma_{z|p}^{-2}\,\big(\mathbb{E}[z \mid rp] - rp\big)^2.$$

Uninformed expected squared Sharpe. Define the *ex-ante* expected squared Sharpe ratio of an uninformed participant as

$$\mu^U(D,\bar{\tau}) \;\equiv\; \mathbb{E}\big[\mathcal{S}_U^2(rp)\big]. \tag{26}$$

For transparency and compactness, we impose the standard normalization $\mu_x = 0$ in what follows (the $\mu_x \neq 0$ case adds only an extra constant-risk-premium term and does not change the key monotonicities in D). Under $\mu_x = 0$, we obtain a closed form:

Proposition 5 (Expected squared Sharpe ratio for an uninformed participant)**.** *Assume* $\mu_x = 0$. *Then*

$$\mu^U(D,\bar{\tau}) = \frac{\sigma_x^6\sigma_z^2}{\Big(\bar{\tau}\big(D^2\sigma_z^2+\sigma_x^2\big) + D\,\sigma_x^2\sigma_z^2\Big)^2}. \tag{27}$$

Moreover, $\partial\mu^U/\partial D < 0$ *and* $\partial\mu^U/\partial\bar{\tau} < 0$.

Proof. With $\mu_x = 0$, $\mu_z - \mathbb{E}[rp] = 0$ in equilibrium (by (25)). Hence $\mathbb{E}[z \mid rp] - rp = (\kappa-1)(rp - \mathbb{E}[rp])$, where κ is defined in (17). Because $\mathbb{V}(z \mid rp) = 1/\sigma_{z|p}^{-2}$ is constant, we have

$$\mu^U = \sigma_{z|p}^{-2}\,\mathbb{E}\big[(\mathbb{E}[z \mid rp] - rp)^2\big] = \sigma_{z|p}^{-2}(\kappa-1)^2\mathbb{V}(rp).$$

Substituting $\mathbb{V}(rp) = b^2\sigma_z^2 + c^2\sigma_x^2$ together with the equilibrium coefficients (24) and $\sigma_{z|p}^{-2} = \sigma_z^{-2} + D^2/\sigma_x^2$ yields (27) after algebra. For monotonicity, write

$$\mu^U = \frac{\sigma_x^6\sigma_z^2}{A(D,\bar{\tau})^2}, \qquad A(D,\bar{\tau}) = \bar{\tau}(D^2\sigma_z^2+\sigma_x^2) + D\sigma_x^2\sigma_z^2.$$

Since $\partial A/\partial D = 2\bar{\tau}D\sigma_z^2 + \sigma_x^2\sigma_z^2 > 0$ and $\partial A/\partial\bar{\tau} = D^2\sigma_z^2 + \sigma_x^2 > 0$, both partial derivatives of μ^U are negative. □

Informed realized Sharpe. An informed participant observes (rp, s_j), so

$$\mathcal{S}_j^2(rp,s_j) = \frac{\big(\mathbb{E}[z \mid rp, s_j] - rp\big)^2}{\mathbb{V}(z \mid rp, s_j)} = \big(\sigma_{z|p}^{-2} + d_j\big)\,\big(\mathbb{E}[z \mid rp, s_j] - rp\big)^2.$$

Informed expected squared Sharpe and the value of private precision. A key object for information demand is how the *expected* squared Sharpe ratio increases with private precision. Let

$$\mu(D,\bar{\tau}) \;\equiv\; \frac{\partial}{\partial\, d_j}\,\mathbb{E}\big[\mathcal{S}_j^2(rp,s_j)\big] \quad \text{(treating investor } j \text{ as infinitesimal)}.$$

In this Gaussian environment, the gain is exactly linear in d_j:

Proposition 6 (Expected Sharpe gain from private precision)**.** *For an infinitesimal investor facing aggregates* $(D, \bar{\tau})$*,*

$$\mathbb{E}\big[\mathcal{S}_j^2(rp, s_j)\big] = \mu^U(D, \bar{\tau}) + \mu(D, \bar{\tau})\, d_j, \tag{28}$$

where

$$\mu(D, \bar{\tau}) = \frac{1 + \mu^U(D, \bar{\tau})}{\sigma_{z|p}^{-2}} = \frac{1 + \mu^U(D, \bar{\tau})}{\sigma_z^{-2} + D^2/\sigma_x^2}. \tag{29}$$

In particular, $\partial\mu/\partial D < 0$ *and* $\partial\mu/\partial\bar{\tau} < 0$*.*

Proof. Fix $(D, \bar{\tau})$, hence fix $\sigma_{z|p}^{-2}$ and the joint law of (z, rp). Write $m(rp) \equiv \mathbb{E}[z \mid rp]$ and $v \equiv \mathbb{V}(z \mid rp) = 1/\sigma_{z|p}^{-2}$. Using (20)–(21),

$$\mathbb{E}[z \mid rp, s_j] - rp = \big(m(rp) - rp\big) + \omega_j\big(s_j - m(rp)\big), \qquad \mathbb{V}(z \mid rp, s_j) = \frac{1}{\sigma_{z|p}^{-2} + d_j}.$$

Conditional on rp, the term $(s_j - m(rp)) = (z - m(rp)) + \varepsilon_j$ has mean 0 and variance $v + \sigma_{s,j}^2$. A short calculation implies

$$\mathbb{E}\big[\mathcal{S}_j^2(rp, s_j) \mid rp\big] = (\sigma_{z|p}^{-2} + d_j)\,(m(rp) - rp)^2 + \frac{d_j}{\sigma_{z|p}^{-2}}.$$

Taking expectations over rp and using $\mu^U = \mathbb{E}[(m(rp) - rp)^2]/v = \sigma_{z|p}^{-2}\mathbb{E}[(m(rp) - rp)^2]$ gives (28) with (29). Monotonicity follows from μ^U decreasing in $(D, \bar{\tau})$ and $\sigma_{z|p}^{-2}$ increasing in D. □

Intuition. μ^U is the expected squared Sharpe ratio available to an *uninformed* (price-only) participant. μ is the *value per unit of private precision*: it is larger when prices are less informative (low $\sigma_{z|p}^{-2}$) and when uninformed opportunities are larger (high μ^U).

A.1.5 Unconditional variance of the risky payoff for uninformed vs. informed investors.

The risky component of investor j's terminal wealth is the payoff from the risky asset position,

$$q_{1j}\,(z - rp).$$

For any information set $\mathcal{I}_j$ at the trading date, the optimal risky demand is

$$q_{1j} = \tau_j\, \frac{\mathbb{E}[z \mid \mathcal{I}_j] - rp}{\mathbb{V}(z \mid \mathcal{I}_j)}.$$

Define the (trading-date) squared Sharpe ratio implied by $\mathcal{I}_j$ as

$$\mathcal{S}_j^2(\mathcal{I}_j) \equiv \frac{(\mathbb{E}[z \mid \mathcal{I}_j] - rp)^2}{\mathbb{V}(z \mid \mathcal{I}_j)}.$$

Then, using the decomposition $z - rp = (\mathbb{E}[z \mid \mathcal{I}_j] - rp) + (z - \mathbb{E}[z \mid \mathcal{I}_j])$ and the fact that $\mathbb{V}(z \mid \mathcal{I}_j)$ does not depend on the realization of signals in the Gaussian environment, we obtain:

$$\mathbb{E}[q_{1j}(z - rp) \mid \mathcal{I}_j] = \tau_j \, \mathcal{S}_j^2(\mathcal{I}_j), \tag{30}$$

$$\mathbb{V}(q_{1j}(z - rp) \mid \mathcal{I}_j) = \tau_j^2 \, \mathcal{S}_j^2(\mathcal{I}_j). \tag{31}$$

Applying the law of total variance yields the unconditional variance:

$$\mathbb{V}\big(q_{1j}(z - rp)\big) = \tau_j^2 \Big(\mathbb{E}[\mathcal{S}_j^2(\mathcal{I}_j)] + \mathbb{V}(\mathcal{S}_j^2(\mathcal{I}_j)) \Big). \tag{32}$$

Under the standard normalization $\mu_x = 0$, we have $\mathbb{E}[rp] = \mathbb{E}[z] = \mu_z$, so $\mathbb{E}[\mathbb{E}[z \mid \mathcal{I}_j] - rp] = 0$. Since $\mathbb{E}[z \mid \mathcal{I}_j] - rp$ is linear in Gaussian signals and $\mathbb{V}(z \mid \mathcal{I}_j)$ is constant, the Sharpe ratio $\mathcal{S}_j(\mathcal{I}_j)$ is mean-zero normal, hence $\mathcal{S}_j^2(\mathcal{I}_j)$ is a scaled $\chi^2(1)$ random variable. Therefore

$$\mathbb{V}\big(\mathcal{S}_j^2(\mathcal{I}_j)\big) = 2 \Big(\mathbb{E}[\mathcal{S}_j^2(\mathcal{I}_j)] \Big)^2. \tag{33}$$

Plugging (33) into (32) gives the useful closed form:

$$\mathbb{V}\big(q_{1j}(z - rp)\big) = \tau_j^2 \Big(\mathbb{E}[\mathcal{S}_j^2(\mathcal{I}_j)] + 2 \big(\mathbb{E}[\mathcal{S}_j^2(\mathcal{I}_j)] \big)^2 \Big). \tag{34}$$

Uninformed (price-only) investor. For an uninformed participant, $\mathcal{I}_U = \sigma(rp)$ and by definition

$$\mathbb{E}[\mathcal{S}_U^2(rp)] = \mu^U(D, \bar{\tau}).$$

Hence, under $\mu_x = 0$,

$$\mathbb{V}\big(q_{1j}^U(z - rp)\big) = \tau_j^2 \Big(\mu^U + 2(\mu^U)^2 \Big). \tag{35}$$

Informed investor. For an informed participant j, $\mathcal{I}_{I,j} = \sigma(rp, s_j)$ and Appendix A.1.4 shows that the expected squared Sharpe ratio is affine in private precision:

$$\mathbb{E}\big[\mathcal{S}_{I,j}^2(rp, s_j)\big] = \mu^U(D, \bar{\tau}) + \mu(D, \bar{\tau}) \, d_j.$$

Therefore, again under $\mu_x = 0$,

$$\mathbb{V}\big(q_{1j}^I(z - rp)\big) = \tau_j^2 \Big(\mu^U + \mu \, d_j + 2 \big(\mu^U + \mu \, d_j \big)^2 \Big). \tag{36}$$

Why informed (and wealthy) investors look riskier to an econometrician. Even though informed investors face a lower *per-share* conditional variance

$$\mathbb{V}(z \mid rp, s_j) = \left(\sigma_{z|p}^{-2} + d_j\right)^{-1} < \mathbb{V}(z \mid rp) = \left(\sigma_{z|p}^{-2}\right)^{-1}, \qquad \sigma_{z|p}^{-2} \equiv \sigma_z^{-2} + D^2/\sigma_x^2,$$

their optimal demand scales positions inversely with perceived variance, so they take larger positions. Unconditionally, the risky payoff variance in (36) exceeds the uninformed variance (35) whenever $d_j > 0$:

$$\mathbb{V}\big(q_{1j}^I(z - rp)\big) - \mathbb{V}\big(q_{1j}^U(z - rp)\big) = \tau_j^2\Big(\mu\, d_j\big(1 + 4\mu^U\big) + 2(\mu\, d_j)^2\Big) > 0.$$

In the baseline (DARA/CRRA-approximation) calibration, $\tau_j = W_{0j}/\tau$, so this risk also scales quadratically with wealth: $\mathbb{V}(q_{1j}(z - rp)) \propto W_{0j}^2$. Thus, wealthy investors who are also better informed can appear to take on more systematic risk in unconditional data, even though their *conditional* risk per share is lower.

A.1.6 Optimal information intensity and wealth cutoffs

Information choice. Combining (14) and Proposition 6, the date-0 expected utility gain from choosing precision d_j (holding aggregates fixed) is

$$\frac{W_{0j}}{2\tau}\left(\mu^U + \mu\, d_j\right) - C(d_j),$$

where μ^U is a constant w.r.t. d_j. Thus the optimal precision solves a quadratic problem with linear benefit:

$$d_j(W_{0j}) = \max\left\{0,\ \frac{\frac{W_{0j}}{2\tau}\mu - c_1}{c_0}\right\} = \max\left\{0,\ \frac{W_{0j}\mu - 2\tau c_1}{2\tau c_0}\right\}. \tag{37}$$

Wealth cutoffs and sorting. An agent enters if the date-0 trading benefit for an uninformed participant covers the entry fee:

$$\frac{W_{0j}}{2\tau}\,\mu^U \ \geq\ F.$$

Hence the participation cutoff is

$$W_0^{\text{part}} = \frac{2\tau F}{\mu^U(D, \bar{\tau})}. \tag{38}$$

Conditional on participation, an agent becomes informed (i.e. chooses $d_j > 0$) iff $\frac{W_{0j}}{2\tau}\mu \geq c_1$, equivalently

$$W_0^{\text{info}} = \frac{2\tau c_1}{\mu(D, \bar{\tau})}. \tag{39}$$

Thus (generically) there are three groups:

$$\begin{aligned} W_{0j} &< W_0^{\text{part}} : \text{ nonparticipant,} \\ W_0^{\text{part}} \le W_{0j} &< W_0^{\text{info}} : \text{ uninformed participant,} \\ W_{0j} &\ge W_0^{\text{info}} : \text{ informed participant (with intensity (37)).} \end{aligned}$$

Remark. Notice that μ^U and μ are not a single object; here they are *distinct* and related by $\mu = (1+\mu^U)/\sigma^{-2}_{z|p}$ (Proposition 6). Participation depends on the *level* of uninformed opportunities μ^U, while information demand depends on the *marginal* value μ.

A.1.7 Population shares: participation, price-only, and informed

Let $G_0(\cdot)$ denote the CDF of initial wealth W_0 in the (unit-mass) household population, with density $g_0(\cdot)$ when it exists. Because entry and information acquisition are monotone in wealth (Section A.1.6), equilibrium sorting is summarized by the two wealth cutoffs (38)–(39):

$$W_0^{\text{part}} = \frac{2\tau F}{\mu^U(D,\bar{\tau})}, \qquad W_0^{\text{info}} = \frac{2\tau c_1}{\mu(D,\bar{\tau})}.$$

Unconditional shares (fractions of the population). Define the participation indicator $\mathbf{1}_{\text{part}}(W_0) \equiv \mathbf{1}\{W_0 \ge W_0^{\text{part}}\}$, and the informed-participation indicator $\mathbf{1}_{\text{info}}(W_0) \equiv \mathbf{1}\{W_0 \ge W_0^{\text{part}}\}\mathbf{1}\{W_0 \ge W_0^{\text{info}}\} = \mathbf{1}\{W_0 \ge \max(W_0^{\text{part}}, W_0^{\text{info}})\}$. Then the equilibrium population shares are:

$$\pi_{\text{part}} \equiv \Pr(W_0 \ge W_0^{\text{part}}) = 1 - G_0\big(W_0^{\text{part}}\big) = \int_{W_0^{\text{part}}}^{\infty} g_0(w)\,dw, \tag{40}$$

$$\pi_{\text{info}} \equiv \Pr\big(W_0 \ge \max(W_0^{\text{part}}, W_0^{\text{info}})\big) = 1 - G_0\big(\max(W_0^{\text{part}}, W_0^{\text{info}})\big) = \int_{\max(W_0^{\text{part}}, W_0^{\text{info}})}^{\infty} g_0(w)\,dw, \tag{41}$$

$$\pi_{\text{U}} \equiv \Pr(\text{price-only participant}) = \Pr\big(W_0 \ge W_0^{\text{part}},\, W_0 < W_0^{\text{info}}\big) = \pi_{\text{part}} - \pi_{\text{info}}. \tag{42}$$

Equivalently, when $W_0^{\text{info}} \ge W_0^{\text{part}}$ (the empirically relevant three-group case),

$$\pi_{\text{U}} = G_0\big(W_0^{\text{info}}\big) - G_0\big(W_0^{\text{part}}\big), \qquad \pi_{\text{info}} = 1 - G_0\big(W_0^{\text{info}}\big), \qquad \pi_{\text{part}} = \pi_{\text{U}} + \pi_{\text{info}}. \tag{43}$$

Finally, the nonparticipation rate is

$$\pi_{\text{N}} \equiv \Pr(W_0 < W_0^{\text{part}}) = G_0\big(W_0^{\text{part}}\big) = 1 - \pi_{\text{part}}. \tag{44}$$

Conditional shares among participants. Sometimes it is useful to report shares *conditional on participation*:

$$\lambda_{\text{info}} \equiv \Pr(\text{informed} \mid \text{participate}) = \frac{\pi_{\text{info}}}{\pi_{\text{part}}}, \qquad \lambda_{\text{U}} \equiv \Pr(\text{price-only} \mid \text{participate}) = \frac{\pi_{\text{U}}}{\pi_{\text{part}}} = 1 - \lambda_{\text{info}}. \tag{45}$$

Connection to the endogenous measure V used in aggregates. If the household population is normalized to unit mass with wealth CDF G_0, the endogenous cross-sectional *measure* over participants used in (22) can be written as the truncated population measure

$$dV(w) = \mathbf{1}\{w \geq W_0^{\text{part}}\}\, dG_0(w), \qquad \int dV = \pi_{\text{part}}.$$

With this convention, $\bar{\tau} = \int \tau(w)\, dV(w)$ and $D = \int \tau(w) d_j(w)\, dV(w)$ are the population aggregates of risk-bearing and (risk-weighted) private precision among participants, respectively.

A.2 Fixed point solution for equilibrium.

Given $(D, \bar{\tau})$, equations (8)–(7) and the policy (10) pin down households' participation and information choices. Aggregating these choices via (4) returns an updated pair $(D, \bar{\tau})$. Equilibrium is therefore a two-dimensional fixed point in $(D, \bar{\tau})$.

A.3 Comparative statics of the wealth thresholds

This section gives clean derivative formulas and sufficient (transparent) conditions under which W_0^{part} and W_0^{info} change when (F, c_1, c_0) change.

Derivative identities. Let equilibrium aggregates $\theta = (D, \bar{\tau})$ depend on a parameter $\xi \in \{F, c_1, c_0\}$. Define $\mu^U(\theta)$ by (26) and $\mu(\theta)$ by (29). From (38)–(39),

$$\frac{dW_0^{\text{part}}}{d\xi} = \frac{2\tau}{\mu^U}\, \mathbf{1}\{\xi = F\} - \frac{2\tau F}{(\mu^U)^2}\, \frac{d\mu^U}{d\xi}, \tag{46}$$

$$\frac{dW_0^{\text{info}}}{d\xi} = \frac{2\tau}{\mu}\, \mathbf{1}\{\xi = c_1\} - \frac{2\tau c_1}{\mu^2}\, \frac{d\mu}{d\xi}. \tag{47}$$

Two monotonicities we can always use. Under $\mu_x = 0$, Proposition 5 and 6 imply:

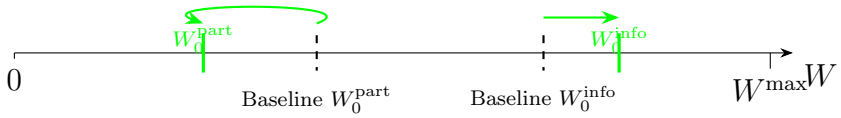
$$\frac{\partial \mu^U}{\partial D} < 0, \quad \frac{\partial \mu^U}{\partial \bar{\tau}} < 0, \qquad \frac{\partial \mu}{\partial D} < 0, \quad \frac{\partial \mu}{\partial \bar{\tau}} < 0. \tag{48}$$

So: (i) more aggregate private information (higher D) reduces both uninformed opportunities and the marginal value of private precision; (ii) more aggregate risk-bearing (higher $\bar{\tau}$) also reduces both.

A.3.1 Lemma 1.1. Entry fee F

If F decreases, then W_0^{part} decreases and W_0^{info} increases, provided the general-equilibrium feedback of μ^U is not too elastic.

Figure 5: **Threshold movements for F cut.** Baseline thresholds W_0^{part} and W_0^{info} (dashed gray) on the wealth axis $[0, W^{\max}]$. When F falls, the wealth thresholds move *further apart* (green).



Proof. Differentiate (38) w.r.t. F:

$$\frac{dW_0^{\text{part}}}{dF} = \frac{2\tau}{\mu^U} - \frac{2\tau F}{(\mu^U)^2}\frac{d\mu^U}{dF} = \frac{2\tau}{\mu^U}\Big(1 - \frac{F}{\mu^U}\frac{d\mu^U}{dF}\Big).$$

Hence $dW_0^{\text{part}}/dF > 0$ (equivalently $F \downarrow \Rightarrow W_0^{\text{part}} \downarrow$) whenever

$$0 \;<\; \frac{F}{\mu^U}\frac{d\mu^U}{dF} \;<\; 1, \qquad \text{i.e.} \qquad 0 \;<\; \frac{d\ln \mu^U}{d\ln F} \;<\; 1. \tag{49}$$

Next, from (47) with $\xi = F$,

$$\frac{dW_0^{\text{info}}}{dF} = -\frac{2\tau c_1}{\mu^2}\frac{d\mu}{dF}.$$

Therefore $F \downarrow \Rightarrow W_0^{\text{info}} \uparrow$ holds whenever $d\mu/dF > 0$. A sufficient (transparent) condition for $d\mu/dF > 0$ is that raising F reduces equilibrium aggregates $(D, \bar{\tau})$ (fewer participants and weaker information acquisition), since μ is decreasing in both components by (48). □

Intuition. Lowering F has a *direct* mechanical effect: it makes entry cheaper, pushing W_0^{part} down. In general equilibrium, more entry deepens risk-sharing and tends to reduce predictable mispricing from noise supply, lowering μ^U; this pushes W_0^{part} *up* and is the reason we state (49) as a sufficient dominance condition. For information, F matters only through equilibrium profitability: with more participation, the marginal value of private precision μ typically falls, raising the wealth threshold W_0^{info} (fewer agents find it worthwhile to become informed).

A.3.2 Lemma 1.2. Linear information cost c_1

If c_1 decreases, then W_0^{info} decreases and W_0^{part} increases, provided the general-equilibrium feedback of μ is not too elastic.

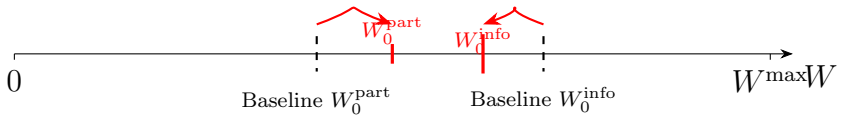


Figure 6: **Threshold movements for c_1 cut.** Baseline thresholds W_0^{part} and W_0^{info} (dashed gray) on the wealth axis $[0, W^{\max}]$. When c_1 falls, the thresholds move *closer together* (red).

Proof. Differentiate (39) w.r.t. c_1:

$$\frac{dW_0^{\text{info}}}{dc_1} = \frac{2\tau}{\mu} - \frac{2\tau c_1}{\mu^2}\frac{d\mu}{dc_1} = \frac{2\tau}{\mu}\Big(1 - \frac{c_1}{\mu}\frac{d\mu}{dc_1}\Big).$$

Thus $dW_0^{\text{info}}/dc_1 > 0$ (equivalently $c_1 \downarrow \Rightarrow W_0^{\text{info}} \downarrow$) whenever

$$0 \;<\; \frac{c_1}{\mu}\frac{d\mu}{dc_1} \;<\; 1, \qquad \text{i.e.} \qquad 0 \;<\; \frac{d\ln\mu}{d\ln c_1} \;<\; 1. \tag{50}$$

For participation,

$$\frac{dW_0^{\text{part}}}{dc_1} = -\frac{2\tau F}{(\mu^U)^2}\frac{d\mu^U}{dc_1}.$$

Hence $c_1 \downarrow \Rightarrow W_0^{\text{part}} \uparrow$ holds whenever $d\mu^U/dc_1 > 0$. A sufficient (transparent) condition for $d\mu^U/dc_1 > 0$ is that raising c_1 reduces equilibrium informativeness (i.e. reduces D), since μ^U is decreasing in D by (48). □

Intuition. Lower c_1 directly reduces the wealth required to justify buying data, so W_0^{info} falls. In general equilibrium, cheaper data increases aggregate private information, making prices more informative and eroding uninformed trading opportunities (lower μ^U). This raises the wealth cutoff for participating without data, W_0^{part}. Condition (50) simply says this feedback does not overturn the direct mechanical effect of a lower c_1 on W_0^{info}.

A.3.3 Lemma 1.3. Quadratic information cost c_0

If c_0 decreases, then W_0^{info} and W_0^{part} move in the same direction. Moreover, if $c_0 \downarrow$ increases equilibrium aggregate information D (the usual case), then both cutoffs increase.

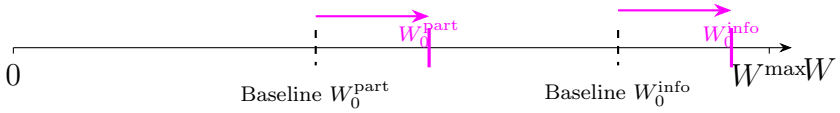


Figure 7: **Threshold movements for c_0 cut.** Baseline thresholds W_0^{part} and W_0^{info} (dashed gray) on the wealth axis $[0, W^{\max}]$. When c_0 falls, the thresholds shift *together to the right* (magenta).

Proof. Since c_0 does not appear directly in (38)–(39),

$$\frac{dW_0^{\text{part}}}{dc_0} = -\frac{2\tau F}{(\mu^U)^2}\frac{d\mu^U}{dc_0}, \qquad \frac{dW_0^{\text{info}}}{dc_0} = -\frac{2\tau c_1}{\mu^2}\frac{d\mu}{dc_0}.$$

Thus W_0^{part} and W_0^{info} move in the same direction whenever $\frac{d\mu^U}{dc_0}$ and $\frac{d\mu}{dc_0}$ share the same sign.

A sufficient condition for this is that c_0 affects equilibrium primarily through aggregate informativeness D and that $dD/dc_0 < 0$ (lower curvature makes informed investors choose higher precision by (37)). Then (48) implies $d\mu^U/dc_0 > 0$ and $d\mu/dc_0 > 0$, so $c_0 \downarrow$ lowers both μ^U and μ, which raises both thresholds. □

Intuition. Lower c_0 steepens the *intensive* response of precision among already-informed investors. That raises aggregate private information and makes prices more informative, which depresses both (i) the expected Sharpe available to an uninformed trader (μ^U) and (ii) the marginal value of adding private precision (μ). Because W_0^{part} and W_0^{info} are both inversely related to these profitability objects, the cutoffs move together.

A.4 What if investors can pool resources and share the cost of data?

The main inequality ranking is robust when investors can delegate information acquisition to a competitive active manager and benefit from pooling. I implement this idea using a parsimonious delegated asset-management setup in the spirit of Gârleanu and Pedersen (2018).[4] The goal is not to build a rich institutional model of intermediation. The goal is to isolate how technology-induced declines in (i) participation/search frictions and (ii) information costs reallocate who enters, who becomes informed, and who benefits.

All institutional details and derivations that map our setup into the competitive delegation frame-

[4]See also Cuoco and Kaniel (2011) and He and Xiong (2013) for related delegated management environments.

work can be accessed on request; the main text focuses on the participation paradox, and on key incidence margins and the resulting inequality ranking.

Each investor chooses among four options: (i) stay out; (ii) enter and trade uninformed; (iii) enter and acquire precision individually (as in the benchmark); or (iv) enter and pay a fixed *search cost* $F_m \geq 0$ to find a manager and join a pooled fund. Managers choose a common signal precision for the pool and charge fees that (by competition) cover information costs.

Pooling creates a simple "club-good" benefit: by sharing the fixed components of information production, delegated learning becomes attractive for investors who would not find it worthwhile to produce information on their own. At the same time, very wealthy (or very risk-tolerant) investors eventually prefer to self-learn because the private benefit of tailoring precision to a large position outweighs the advantages of pooling. This generates a *delegation window*: a range of investors for whom delegation dominates trading uninformed, but not necessarily self-learning.

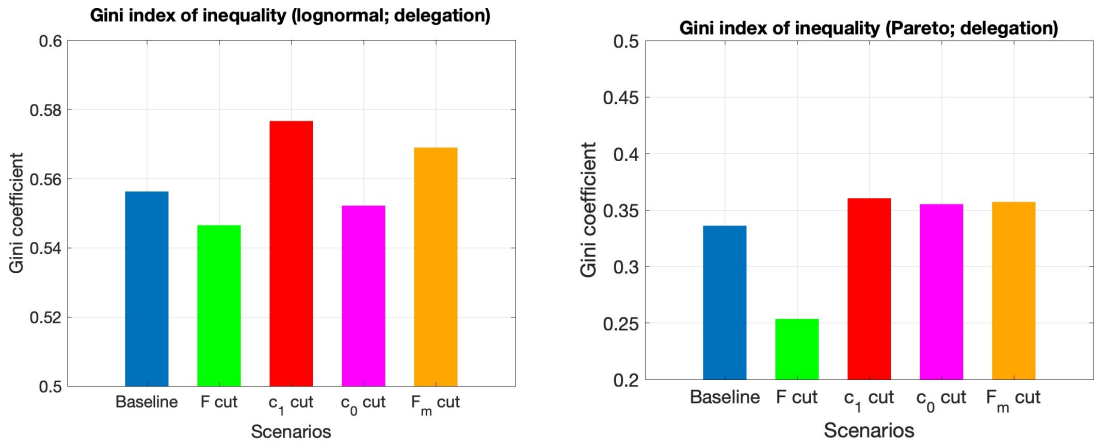
Delegation changes *how* information is purchased, but it does not change the model's basic price-formation logic. The delegated sector adds to aggregate private precision D, so a larger delegated sector makes prices more informative. Compared to the benchmark, there are now three cutoffs: a participation cutoff W^{part}, a delegation cutoff W^{del} above which investors prefer to delegate rather than remain uninformed, and a switch cutoff W^{switch} above which investors prefer to self-learn rather than delegate.

With delegation, technological improvements still map into inequality through the same three intuitive margins: (i) who enters at the participation cutoff, (ii) who moves from uninformed participation into (delegated) information, and (iii) who moves from delegation into self-learning at the top.

The overall effect on the Lorenz curve or the Gini depends on how much mass moves at each cutoff (which is determined by the wealth density) and on the general-equilibrium feedback through $(D, \bar{\tau})$, which shifts the marginal value of information μ. Because both the masses and the feedback are endogenous, we do not claim a universal ranking between delegation and self-learning in all parameterizations.

The qualitative ranking mirrors the benchmark. Lower participation costs ($F \downarrow$) expand entry at the left margin and typically compress inequality (a between-group compression effect). Lower information adoption costs ($c_1 \downarrow$) increase the mass of informed investors (delegated and self), thicken the right tail among participants, and raise inequality among participants. Lower information intensity costs ($c_0 \downarrow$) mainly scales up precision for agents who are already informed (delegated and self), so its aggregate inequality effects are typically smaller and calibration-dependent. Lower manager search

Figure 8: **Inequality across scenarios (delegation extension).** Left: Lognormal prior. Right: Pareto prior. Gini coefficients of final wealth W_2 in the Baseline (blue) and after one-at-a-time 70% cuts in the participation cost F (green), and a 50% cut in the marginal information cost c_1 (red), the quadratic information cost c_0 (magenta), and the search for a manager cost F_m (orange). Cutting F compresses the distribution and lowers inequality; cutting c_1 widens it and raises inequality; cutting c_0 and F_m leave inequality nearly unchanged. Colors are consistent across panels.



Notes. Baseline (blue): lognormal (left) and Pareto (right) W_0; $\rho = 0.90$, $h_0 = 0.1$, $\sigma_x^2 = 1$, $(F, c_1, c_0) = (2, 2, 2)$, $Fm = 50$, bond gross return $r_{\text{bond}} = 1.07$. Calibration targets: participation 60% and self-informed 3.4%. For each scenario we solve the fixed-point conditions for (s, t, T_m, f, h_j, h_m), compute W_{part}, W_{del}, W_{info} and W_{switch}, map W_0 into W_2, and then compute the Gini and the Lorenz curve. Reported Ginis with lognormal prior: baseline 0.556 (blue), F-cut 0.547 (green), c_1-cut 0.577 (red), c_0-cut 0.525 (magenta), F_m-cut 0.556 (orange). Reported Ginis with Pareto prior: baseline 0.336 (blue), F-cut 0.254 (green), c_1-cut 0.360 (red), c_0-cut 0.355 (magenta), F_m-cut 0.336 (orange). Variable definitions are in Appendix **??**.

costs ($F_m \downarrow$) expand delegation by making it cheaper to access pooled information. This pushes some uninformed participants into delegation, but it can also reduce entry indirectly by changing equilibrium informativeness and thus the value of entering. Which force dominates is theoretically ambiguous and is therefore quantified in the calibration.

Figure 8 reports how inequality responds to each cost reduction in the delegation extension under the lognormal prior (left) and Pareto prior (right). Across priors, the *ranking* above is robust. The magnitude of delegation-driven changes is modest relative to the benchmark, and the main movements occur in the right tail when information frictions fall.

A full analytical derivation is available on request.

A.5 A Dynamic Extension with Participation Persistence, Reentry Costs, Labor Income, Housing, and Experience-Based Beliefs

This subsection extends the static participation–information problem to a finite-horizon dynamic environment while preserving the paper's one-period Gaussian information block. The objective is to show that the adverse-selection mechanism survives once participation is persistent, reentry is costly, labor income and housing affect effective liquidity, and households update beliefs after bad aggregate market outcomes.

State variables and status transitions. Time is indexed by $t = 0, 1, \ldots, T-1$. The household state is

$$S_t = (x_t, \ell_t, h_t, b_t, \sigma_t),$$

where $x_t > 0$ is effective liquid wealth, ℓ_t is a labor-income state, h_t is a housing state, $b_t \geq 0$ is a stock of bad market experience, and $\sigma_t \in \{N, R, A\}$ is participation status. The three status values have the interpretation

$$N = \text{never entered}, \qquad R = \text{inactive but previously entered}, \qquad A = \text{active last period}.$$

The participation cost depends on status:

$$K(\sigma_t) = \begin{cases} F_E, & \sigma_t = N, \\ F_R, & \sigma_t = R, \\ f, & \sigma_t = A, \end{cases} \qquad \text{with } f \leq F_R \leq F_E.$$

If the household participates at date t, then next period's status is $\sigma_{t+1} = A$. If it does not participate, then

$$\sigma_{t+1} = \begin{cases} N, & \sigma_t = N, \\ R, & \sigma_t \in \{R, A\}. \end{cases}$$

Thus, staying active is cheaper than reentering, and reentering is cheaper than first entry.

We define effective liquid wealth as

$$x_t \equiv a_t + \omega_Y Y(\ell_t) - \omega_H H(h_t),$$

where a_t is liquid financial wealth, $Y(\ell_t)$ is current labor income (or the liquid component of human wealth), $H(h_t)$ is the household's housing commitment, and $\omega_Y, \omega_H \in [0,1]$ convert those objects into resources that are relevant for direct risky-market participation. In this appendix extension, (x_t, ℓ_t, h_t, b_t) evolve as exogenous Markov states from the viewpoint of the within-period participation and information problem; only the status component σ_t is action-dependent. This is the minimal dynamic extension needed to study persistence and reentry analytically.

The bad-experience state evolves according to

$$b_{t+1} = \rho_b b_t + (1-\rho_b)\max\{-r^m_{t+1}, 0\}, \qquad 0 \le \rho_b < 1,$$

where r^m_{t+1} is the aggregate market excess return. Let

$$\delta_t \equiv \delta(b_t) = e^{-\eta b_t}, \qquad \eta > 0,$$

denote the experience-based attenuation of perceived risky-return opportunities. After bad aggregate outcomes, b_t rises and δ_t falls.

Within-period trading block. Let the risky asset's payoff be $z_{t+1} \sim N(\mu_z, \sigma_z^2)$ and let noise-trader supply have variance σ_x^2, exactly as in the static model. Let $(D_t, \bar{\tau}_t)$ denote the aggregate risk-tolerance-weighted private precision and aggregate risk-bearing capacity at date t. As in the baseline model, the price reveals information through the posterior precision

$$\lambda_{p,t} \equiv \sigma^{-2}_{z|p,t} = \sigma_z^{-2} + \frac{D_t^2}{\sigma_x^2}.$$

Define the price-only expected excess payoff as

$$a_t(p_t) \equiv \mathbb{E}_t[z_{t+1} \mid p_t] - r p_t.$$

A participating household can buy a private signal

$$s_t = z_{t+1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, d_t^{-1}),$$

with precision $d_t \ge 0$ and cost

$$C(d_t) = c_1 d_t + \frac{c_0}{2} d_t^2, \qquad c_1 > 0,\ c_0 > 0.$$

Conditional on (p_t, s_t), the posterior mean and variance are

$$\mathbb{E}_t[z_{t+1} \mid p_t, s_t] = \mathbb{E}_t[z_{t+1} \mid p_t] + \frac{d_t}{\lambda_{p,t} + d_t}\Big(s_t - \mathbb{E}_t[z_{t+1} \mid p_t]\Big),$$

$$\mathbb{V}_t(z_{t+1} \mid p_t, s_t) = \frac{1}{\lambda_{p,t} + d_t}.$$

To keep the dynamic formulas transparent, labor income risk and housing affect participation through effective liquidity x_t and the continuation value, rather than through a direct hedge term in contemporaneous risky demand. Under that orthogonality restriction, the certainty-equivalent gain from trading with information set I_t is

$$\mathcal{G}_t(I_t) = \frac{x_t}{2\tau}\,\widetilde{S}_t(I_t)^2, \qquad \widetilde{S}_t(I_t)^2 \equiv \delta_t^2 \frac{\mathbb{E}_t[z_{t+1} - rp_t \mid I_t]^2}{\mathbb{V}_t(z_{t+1} - rp_t \mid I_t)}.$$

Thus, bad experiences scale down the risky-return opportunities that the household perceives.

Dynamic analogues of μ_U and μ. Fix $(D_t, \bar{\tau}_t)$ and treat the household as infinitesimal. Define the dynamic price-only opportunity set

$$\widetilde{\mu}_{U,t}(b_t; D_t, \bar{\tau}_t) \equiv \mathbb{E}_t\Big[\widetilde{S}_t(p_t)^2\Big].$$

Since $\widetilde{S}_t(p_t)^2 = \delta_t^2 S_t(p_t)^2$, where $S_t(p_t)^2$ is the static price-only squared Sharpe ratio, we obtain immediately

$$\widetilde{\mu}_{U,t}(b_t; D_t, \bar{\tau}_t) = \delta_t^2 \mu_U(D_t, \bar{\tau}_t).$$

Likewise, for an informed household with precision choice d_t,

$$\mathbb{E}_t\Big[\widetilde{S}_t(p_t, s_t)^2\Big] = \delta_t^2\Big(\mu_U(D_t, \bar{\tau}_t) + \mu(D_t, \bar{\tau}_t) d_t\Big).$$

Therefore the dynamic marginal value of private precision is

$$\widetilde{\mu}_t(b_t; D_t, \bar{\tau}_t) = \delta_t^2 \mu(D_t, \bar{\tau}_t).$$

These equalities are exact in the orthogonal-background-risk case. They show that experience-based pessimism scales down both the value of entering as a price-only trader and the value of becoming informed, but it does not alter the sign of the comparative statics with respect to aggregate informativeness D_t.

Bellman equation. Let $V_t(S_t)$ denote the household's continuation value. The value of not participating is

$$V_t^0(S_t) = \beta\,\mathbb{E}_t\Big[V_{t+1}(x_{t+1}, \ell_{t+1}, h_{t+1}, b_{t+1}, \sigma_t^0)\Big],$$

where

$$\sigma_t^0 = \begin{cases} N, & \sigma_t = N, \\ R, & \sigma_t \in \{R, A\}. \end{cases}$$

The value of participating is

$$V_t^1(S_t) = -K(\sigma_t) + \max_{d_t \geq 0} \left\{ \frac{x_t}{2\tau} \left(\widetilde{\mu}_{U,t} + \widetilde{\mu}_t d_t \right) - c_1 d_t - \frac{c_0}{2} d_t^2 + \beta \, \mathbb{E}_t \Big[V_{t+1}(x_{t+1}, \ell_{t+1}, h_{t+1}, b_{t+1}, A) \Big] \right\}.$$

The Bellman equation is

$$V_t(S_t) = \max \big\{ V_t^0(S_t), V_t^1(S_t) \big\}.$$

Optimal precision and participation surplus. Because the one-period benefit is affine in d_t and the cost is quadratic, the optimal precision choice is

$$d_t^*(S_t) = \frac{1}{c_0} \left[\frac{x_t \widetilde{\mu}_t}{2\tau} - c_1 \right]_+ = \frac{1}{c_0} \left[\frac{x_t \delta_t^2 \mu(D_t, \bar{\tau}_t)}{2\tau} - c_1 \right]_+ .$$

Define the continuation-value gain from being active next period as

$$\Omega_t(x_t, \ell_t, h_t, b_t, \sigma_t) \equiv \mathbb{E}_t \Big[V_{t+1}(x_{t+1}, \ell_{t+1}, h_{t+1}, b_{t+1}, A) - V_{t+1}(x_{t+1}, \ell_{t+1}, h_{t+1}, b_{t+1}, \sigma_t^0) \Big].$$

Substituting the optimal d_t^* into $V_t^1 - V_t^0$ yields the dynamic participation surplus

$$\Delta_t(S_t) \equiv V_t^1(S_t) - V_t^0(S_t) = -K(\sigma_t) + \frac{x_t \delta_t^2 \mu_U(D_t, \bar{\tau}_t)}{2\tau} + \frac{1}{2c_0} \left[\frac{x_t \delta_t^2 \mu(D_t, \bar{\tau}_t)}{2\tau} - c_1 \right]_+^2 + \beta \, \Omega_t(x_t, \ell_t, h_t, b_t, \sigma_t).$$

The household participates if and only if $\Delta_t(S_t) \geq 0$.

Dynamic participation cutoff. For fixed $(\ell_t, h_t, b_t, \sigma_t, D_t, \bar{\tau}_t)$, the surplus $\Delta_t(S_t)$ is weakly increasing in x_t and strictly increasing whenever either $\mu_U(D_t, \bar{\tau}_t) > 0$ or the information margin is interior. Hence there exists a unique participation cutoff

$$X_t^{\text{part}}(\ell_t, h_t, b_t, \sigma_t; D_t, \bar{\tau}_t)$$

that solves

$$-K(\sigma_t) + \frac{X_t^{\text{part}} \delta_t^2 \mu_U(D_t, \bar{\tau}_t)}{2\tau} + \frac{1}{2c_0} \left[\frac{X_t^{\text{part}} \delta_t^2 \mu(D_t, \bar{\tau}_t)}{2\tau} - c_1 \right]_+^2 + \beta \, \Omega_t(X_t^{\text{part}}, \ell_t, h_t, b_t, \sigma_t) = 0.$$

In the empirically relevant region where the marginal entrant is still uninformed,

$$\frac{x_t \delta_t^2 \mu(D_t, \bar{\tau}_t)}{2\tau} \leq c_1,$$

so the information term is zero and the cutoff simplifies to

$$X_t^{\text{part}}(\ell_t, h_t, b_t, \sigma_t; D_t, \bar{\tau}_t) = \frac{2\tau\,(K(\sigma_t) - \beta\,\Omega_t(\ell_t, h_t, b_t, \sigma_t))}{\delta_t^2 \mu_U(D_t, \bar{\tau}_t)}.$$

This expression displays directly how the dynamic participation threshold depends on persistence and reentry costs through $K(\sigma_t)$, on labor-income and housing conditions through the continuation value Ω_t, on bad experiences through δ_t, and on adverse selection through $\mu_U(D_t, \bar{\tau}_t)$.

Why adverse selection survives. The paper proves that in the static equilibrium,

$$\frac{\partial \mu_U(D_t, \bar{\tau}_t)}{\partial D_t} < 0, \qquad \frac{\partial \mu(D_t, \bar{\tau}_t)}{\partial D_t} < 0.$$

Because the dynamic counterparts are just δ_t^2 multiples of these objects, the same monotonicity carries over:

$$\frac{\partial \widetilde{\mu}_{U,t}}{\partial D_t} < 0, \qquad \frac{\partial \widetilde{\mu}_t}{\partial D_t} < 0.$$

Holding the transition fixed, the derivative of the participation surplus with respect to current aggregate informativeness is therefore

$$\frac{\partial \Delta_t(S_t)}{\partial D_t} = \frac{x_t \delta_t^2}{2\tau} \frac{\partial \mu_U(D_t, \bar{\tau}_t)}{\partial D_t} + \frac{x_t \delta_t^2}{2\tau c_0} \left[\frac{x_t \delta_t^2 \mu(D_t, \bar{\tau}_t)}{2\tau} - c_1 \right]_+ \frac{\partial \mu(D_t, \bar{\tau}_t)}{\partial D_t} < 0.$$

Thus, a rise in D_t lowers the dynamic participation surplus state by state. In the price-only region, the result is immediate:

$$\frac{\partial \Delta_t(S_t)}{\partial D_t} = \frac{x_t \delta_t^2}{2\tau} \frac{\partial \mu_U(D_t, \bar{\tau}_t)}{\partial D_t} < 0.$$

Equivalently, the price-only participation cutoff rises with aggregate informativeness:

$$\frac{\partial X_t^{\text{part}}}{\partial D_t} > 0.$$

If aggregate informativeness is persistent so that larger D_t also predicts larger future D_{t+s}, then the continuation term Ω_t will typically reinforce this result by making future direct participation less attractive as well.

Implication. The dynamic extension preserves the logic of the baseline model. Declines in data costs raise the information intensity of the informed sector, which raises D_t. A larger D_t makes prices more informative, compresses the risk-adjusted opportunity set of price-only households, and increases the effective wealth threshold for direct participation. Persistence and reentry costs amplify the extensive-margin effect because, once a household exits, getting back into direct risky-asset markets is more expensive than simply remaining active.

A.5.1 Mapping observables into participation

This subsection provides a compact mapping from observable return moments and group shares into the model's participation cutoff and participation rate. The key object is the expected squared Sharpe ratio available to a *price-only* (uninformed) participant, μ^U, which is directly identified from return predictability.

Step 1: Identify μ^U from return data. Let y denote the risky asset's *excess payoff in date-2 units*,

$$y \;\equiv\; z - rp,$$

(or, empirically, an excess *return* at the corresponding horizon; since Sharpe ratios are scale-free, the mapping below applies to either payoffs or returns).

For a price-only participant with information set $\mathcal{I}_U = \sigma(rp)$, the (date-1) squared Sharpe ratio is

$$\mathcal{S}_U^2(rp) \;\equiv\; \frac{\mathbb{E}[y \mid rp]^2}{\mathbb{V}(y \mid rp)} = \frac{(\mathbb{E}[z \mid rp] - rp)^2}{\mathbb{V}(z \mid rp)}.$$

Because (y, rp) is jointly normal in a linear equilibrium, $\mathbb{V}(y \mid rp)$ is constant (does not depend on rp). Hence the planning-date expected squared Sharpe ratio is

$$\mu^U \;\equiv\; \mathbb{E}\big[\mathcal{S}_U^2(rp)\big] = \frac{\mathbb{E}\big[\mathbb{E}[y \mid rp]^2\big]}{\mathbb{V}(y \mid rp)} = \frac{\mathbb{E}[y]^2 + \mathbb{V}\big(\mathbb{E}[y \mid rp]\big)}{\mathbb{V}(y \mid rp)}. \tag{51}$$

Using the law of total variance,

$$\mathbb{V}(y) = \mathbb{V}\big(\mathbb{E}[y \mid rp]\big) + \mathbb{V}(y \mid rp),$$

define the (population) predictability statistic

$$R_p^2 \;\equiv\; \frac{\mathbb{V}\big(\mathbb{E}[y \mid rp]\big)}{\mathbb{V}(y)}. \tag{52}$$

In the Gaussian environment, $\mathbb{E}[y \mid rp]$ is the linear projection of y on rp, so R_p^2 equals the population R^2 from the predictive regression of y on rp (or any equivalent one-dimensional sufficient statistic extracted from price). Then $\mathbb{V}(y \mid rp) = \mathbb{V}(y)\,(1 - R_p^2)$ and (51) becomes

$$\boxed{\mu^U = \frac{\underbrace{\dfrac{\mathbb{E}[y]^2}{\mathbb{V}(y)}}_{\text{unconditional squared Sharpe } S_0^2} + R_p^2}{1 - R_p^2}.} \tag{53}$$

Under the normalization $\mu_x = 0$ (zero net supply), $\mathbb{E}[y] = 0$ and this simplifies further to

$$\mu^U = \frac{R_p^2}{1 - R_p^2}.$$

Equation (53) is convenient for quantification because it identifies μ^U from observable return moments and a single predictability statistic R_p^2.

Step 2: Map μ^U into the participation cutoff and participation rate. In the model, a household enters as an uninformed trader iff its expected trading gain covers the fixed entry cost:

$$\frac{W_0}{2\rho}\,\mu^U \;\geq\; F.$$

Thus the participation cutoff is

$$\boxed{W_0^{\text{part}} = \frac{2\rho F}{\mu^U}.} \tag{54}$$

Let G_0 denote the CDF of initial wealth W_0 in the population. The model-implied participation rate is

$$\boxed{\pi_{\text{part}} = \Pr(W_0 \geq W_0^{\text{part}}) = 1 - G_0\left(\frac{2\rho F}{\mu^U}\right).} \tag{55}$$

Combining (53) and (55) yields a direct mapping from observables $(\mathbb{E}[y], \mathbb{V}(y), R_p^2)$ into participation, up to the single composite parameter $2\rho F$.

Step 3: Invert for the entry-cost composite from observed group shares. Let π_N denote the observed nonparticipation rate, so $\pi_N = 1 - \pi_{\text{part}}$. Then W_0^{part} is the π_N-quantile of the wealth distribution:

$$W_0^{\text{part}} = G_0^{-1}(\pi_N).$$

Using (54), the entry-cost composite is identified by

$$\boxed{2\rho F = \mu^U\, G_0^{-1}(\pi_N).} \tag{56}$$

This inversion is useful for calibration: given $(\mathbb{E}[y], \mathbb{V}(y), R_p^2)$ one obtains μ^U from (53), and given the observed nonparticipation rate π_N one obtains $2\rho F$ from (56). Counterfactual changes in μ^U (driven by technology through price informativeness) then map into counterfactual participation rates via (55).

Group shares. Given observed fractions of nonparticipants, price-only participants, and informed participants (π_N, π_U, π_I) with $\pi_N + \pi_U + \pi_I = 1$, the model's participation rate is $\pi_{\text{part}} = \pi_U + \pi_I$. The mapping above explains how π_{part} varies with the return opportunities available to a price-only

trader, summarized by μ^U.