Investor Sophistication and Capital Income Inequality ONLINE APPENDIX

Marcin Kacperczyk
Imperial College

Jaromir Nosal
Boston College

Luminita Stevens
University of Maryland

Abstract

This file contains supplementary material for the paper 'Investor Sophistication and Capital Income Inequality', by Marcin Kacperczyk, Jaromir Nosal, and Luminita Stevens.

Contents

List of Figures

1 Appendix: Proofs

1.1 Model

Portfolio Choice. In the second stage, each investor chooses portfolio holdings q_{ji} to solve

$\max_{\{q_{ji}\}_{i=1}^n} U_j = E_j(W_j) - \frac{\rho}{2} V_j(W_j) \quad s.t. \;\; W_j = r\left(W_{0j} - \sum_{i=1}^n q_{ji} p_i\right) + \sum_{i=1}^n q_{ji} z_i,$

where E_j and V_j denote the mean and variance conditional on investor j's information set:

$E_j(W_j) = E_j\left[rW_{0j} + \sum_{i=1}^n q_{ji}(z_i - rp_i)\right] = rW_{0j} + \sum_{i=1}^n q_{ji}\left[E_j(z_i) - rp_i\right],$

$V_j(W_j) = V_j\left[rW_{0j} + \sum_{i=1}^n q_{ji}(z_i - rp_i)\right] = \sum_{i=1}^n q_{ji}^2 V_j(z_i).$

Let $\widehat{\mu}_{ji} \equiv E_j[z_i]$ and $\widehat{\sigma}_{ji}^2 \equiv V_j[z_i]$. The investor's portfolio problem is to maximize

$U_j = rW_{0j} + \sum_{i=1}^n q_{ji}(\widehat{\mu}_{ji} - rp_i) - \frac{\rho}{2} \sum_{i=1}^n q_{ji}^2 \widehat{\sigma}_{ji}^2.$

The first order conditions with respect to q_{ji} yield $q_{ji} = \frac{\widehat{\mu}_{ji} - rp_i}{\rho \widehat{\sigma}_{ji}^2}$. Since W_{0j} does not affect the optimization, we normalize it to zero. The indirect utility function becomes

$U_j = \frac{1}{2\rho} \sum_{i=1}^n \frac{(\widehat{\mu}_{ji} - rp_i)^2}{\widehat{\sigma}_{ji}^2}.$ $\square$

Posterior Beliefs. The signal structure, $z_i = s_{ji} + \delta_{ji}$, implies that

$\widehat{\mu}_{ji} = \overline{z} + \frac{Cov(s_{ji}, z_i)}{\sigma_{sji}^2}(s_{ji} - \overline{s}_{ji}) = s_{ji},$

$\widehat{\sigma}_{ji}^2 = \sigma_i^2 \left(1 - \frac{Cov^2(s_{ji}, z_i)}{\sigma_{sji}^2 \sigma_i^2}\right) = \sigma_{\delta ji}^2.$ $\square$

Information Constraint. Let $H(z)$ denote the entropy of z, and let $H(z|s_j)$ denote the conditional entropy of z given the vector of signals s_j. Then

$I(z; s_j) \equiv H(z) - H(z|s_j) \overset{(1)}{=} \sum_{i=1}^n H(z_i) - H(z|s_j) \overset{(2)}{=} \sum_{i=1}^n H(z_i) - \sum_{i=1}^n H(z_i|z^{i-1}, s_j)$

$$\overset{(1)}{=} \sum_{i=1}^{n} H(z_i) - \sum_{i=1}^{n} H(z_i|s_j) \overset{(3)}{=} \sum_{i=1}^{n} H(z_i) - \sum_{i=1}^{n} H(z_i|s_{ji}) = \sum_{i=1}^{n} I(z_i; s_{ji})$$

where (1) follows from the independence of the payoffs z_i; (2) follows from the chain rule for entropy, where $z^{i-1} = \{z_1, ..., z_{i-1}\}$; (3) follows from the independence of the signals s_{ji}.

For each asset i, the entropy of $z_i \sim \mathcal{N}(\overline{z}, \sigma_i^2)$ is $H(z_i) = \frac{1}{2}\ln(2\pi e \sigma_i^2)$.

The signal structure, $z_i = s_{ji} + \delta_{ji}$, implies that

$$I(z_i; s_{ji}) = H(z_i) + H(s_{ji}) - H(z_i, s_{ji}) = \frac{1}{2}\log\left(\frac{\sigma_i^2\sigma_{sji}^2}{\left|\Sigma_{z_i s_{ji}}\right|}\right) = \frac{1}{2}\log\left(\frac{\sigma_i^2}{\sigma_{\delta ji}^2}\right),$$

where $\left|\Sigma_{z_i s_{ji}}\right| = \sigma_{sji}^2\sigma_{\delta ji}^2$ is the determinant of the variance-covariance matrix of z_i and s_{ji}.

Hence $I(z_i; s_{ji}) = 0$ if $\sigma_{\delta ji}^2 = \sigma_i^2$.

Across assets, $I(z; s_j) = \sum_{i=1}^{n} I(z_i; s_{ji}) = \frac{1}{2}\sum_{i=1}^{n}\log\left(\frac{\sigma_i^2}{\sigma_{\delta ji}^2}\right) = \frac{1}{2}\log\left(\prod_{i=1}^{n}\frac{\sigma_i^2}{\sigma_{\delta ji}^2}\right) \leq K_j.$

$\square$

Information Objective. Expected utility is given by

$$E_{0j}[U_j] = \frac{1}{2\rho}E_{0j}\left[\sum_{i=1}^{n}\frac{(\widehat{\mu}_{ji} - rp_i)^2}{\widehat{\sigma}_{ji}^2}\right] = \frac{1}{2\rho}\sum_{i=1}^{n}\frac{E_{0j}\left[(\widehat{\mu}_{ji} - rp_i)^2\right]}{\widehat{\sigma}_{ji}^2} = \frac{1}{2\rho}\sum_{i=1}^{n}\left(\frac{\widehat{R}_{ji}^2 + \widehat{V}_{ji}}{\widehat{\sigma}_{ji}^2}\right),$$

where $\widehat{R}_{ji}$ and $\widehat{V}_{ji}$ denote the ex-ante mean and variance of expected excess returns, $\widehat{\mu}_{ji} - rp_i$. Conjecture (and later verify) that prices are normally distributed, $p_i \sim \mathcal{N}\left(\overline{p}_i, \sigma_{pi}^2\right)$.

$$\widehat{R}_{ji} \equiv E_{0j}(\widehat{\mu}_{ji} - rp_i) = \overline{z} - r\overline{p}_i,$$

$$\widehat{V}_{ji} \equiv V_{0j}(\widehat{\mu}_{ji} - rp_i) = Var(\widehat{\mu}_{ji}) + r^2\sigma_{pi}^2 - 2rCov(\widehat{\mu}_{ji}, p_i).$$

The signal structure implies that $Var(\widehat{\mu}_{ji}) = \sigma_{sji}^2$.

Following Admati (1985), we conjecture (and later verify) that prices are $p_i = a_i + b_i\varepsilon_i - c_i\nu_i$, for some coefficients $a_i, b_i, c_i \geq 0$. We compute $Cov(\widehat{\mu}_{ji}, p_i)$ exploiting the fact that posterior beliefs and prices are conditionally independent given payoffs. We obtain

$$\widehat{V}_{ji} = \sigma_{sji}^2 + r^2\sigma_{pi}^2 - 2rb_i\sigma_{sji}^2 = (1 - rb_i)^2\sigma_i^2 + r^2c_i^2\sigma_x^2 - (1 - 2rb_i)\widehat{\sigma}_{ji}^2.$$

Hence the distribution of expected excess returns is normal with mean and variance:

$$\widehat{R}_{ji} = \overline{z} - ra_i \quad \text{and} \quad \widehat{V}_{ji} = (1 - rb_i)^2 \sigma_i^2 + r^2 c_i^2 \sigma_x^2 - (1 - 2rb_i)\, \widehat{\sigma}_{ji}^2.$$

Expected utility becomes

$$E_{0j}[U_j] = \tfrac{1}{2\rho} \textstyle\sum_{i=1}^n G_i \frac{\sigma_i^2}{\widehat{\sigma}_{ji}^2} - \tfrac{1}{2\rho} \sum_{i=1}^n (1 - 2rb_i),$$

where $G_i \equiv (1 - rb_i)^2 + \frac{r^2 c_i^2 \sigma_x^2}{\sigma_i^2} + \frac{(\overline{z} - ra_i)^2}{\sigma_i^2}$, and where the second summation is independent of the investor's choices. $\square$

Proof of Lemma 1 (Information Choice). The linear objective function and the convex constraint imply that each investor allocates all capacity to learning about a single asset. Let $\widehat{\sigma}_{ji}^2 = e^{-2K_{ji}} \sigma_i^2$. Then the optimization problem can be written as $\max_{\{K_{ji}\}_{i=1}^n} \sum_{i=1}^n G_i e^{2K_{ji}} \;\; s.t. \;\; \sum_{i=1}^n K_{ji} \leq K_j$. Suppose the investor allocates capacity to learning about 2 assets. WLOG, let these assets be indexed by 1 and 2, and suppose $G_2 \leq G_1$. Then, $K_{j1} + K_{j2} = K_j$ and the value of the objective function is

$$\sum_{i=1}^n G_i e^{2K_{ji}} = G_1 \left(e^{2K_{j1}} - 1\right) + G_2 \left(e^{2K_{j2}} - 1\right) + \sum_{i=1}^n G_i$$

$$\overset{(1)}{\leq} G_1 \left(e^{2K_{j1}} - 1\right) + G_1 \left(e^{2K_{j2}} - 1\right) + \sum_{i=1}^n G_i \overset{(2)}{<} G_1 \left(e^{2K_j} - 1\right) + \sum_{i=1}^n G_i,$$

where (1) follows from the assumption WLOG that $G_2 \leq G_1$, (2) follows from the fact that for $K_{j1} + K_{j2} = K_j$, $e^{2K_{j1}} + e^{2K_{j2}}$ is minimized at $K_{j1} = K_{j2} = K_j/2$,, strictly convex and maximized at a corner. This chain of inequalities shows that splitting capacity across two assets yields strictly lower utility than investing all capacity in a single asset, even if the gains from learning are equal across assets. Splitting capacity among more than two assets would lower utility even more.

Let l_j index the asset about which investor j learns. The investor's objective is $\left(e^{2K_j} - 1\right) G_{l_j} + \sum_{i=1}^n G_i$. Since $e^{2K_j} > 1$, the objective is maximized by allocating all capacity to the asset with the largest utility gain: $l_j \in \arg\max_i G_i$. The variance of the investor's posterior beliefs is

$$\widehat{\sigma}_{ji}^{2} = \begin{cases} e^{-2K_j}\sigma_i^2 & \text{if } i = l_j, \\ \sigma_i^2 & \text{if } i \neq l_j. \end{cases}$$

$\square$

Conditional Distribution of Signals. Conditional on the realized payoff, the signal is a normally distributed random variable, with mean and variance given by

$$E\left(s_{ji}|z_i\right) = \overline{s}_{ji} + \frac{Cov(s_{ji}, z_i)}{\sigma_i^2}\left(z_i - \overline{z}\right) = \begin{cases} \overline{z} + \left(1 - e^{-2K_j}\right)\varepsilon_i & \text{if } i = l_j \\ \overline{z} & \text{if } i \neq l_j, \end{cases}$$

$$V\left(s_{ji}|z_i\right) = \sigma_{sji}^2\left(1 - \frac{Cov^2(s_{ji}, z_i)}{\sigma_{sji}^2\sigma_i^2}\right) = \begin{cases} \left(1 - e^{-2K_j}\right)e^{-2K_j}\sigma_i^2 & \text{if } i = l_j \\ 0 & \text{if } i \neq l_j. \end{cases} \quad \square$$

Holdings. Individual portfolio holdings are given by

$$q_{ji} = \frac{\widehat{\mu}_{ji} - rp_i}{\rho\widehat{\sigma}_{ji}^2} = \frac{e^{2K_{ji}}(\widehat{\mu}_{ji} - rp_i)}{\rho\sigma_i^2}.$$

For the assets that investor j does not learn about, $K_{ji} = 0$, and $q_{ji} = \left(\overline{z} - rp_i\right)/\rho\sigma_i^2$. This quantity is the same for all investors, regardless of their type.

For the asset that investor j learns about, $K_{ji} = K_j$, and $q_{ji} = e^{2K_j}\left(\widehat{\mu} - rp_i\right)/\rho\sigma_i^2$.

Total holdings for the group of sophisticated investors learning about this asset are

$$\int_{M_{1i}} q_{ji}dj = \int_{M_{1i}} \frac{e^{2K_1}(\widehat{\mu}_{ji} - rp_i)}{\rho\sigma_i^2}dj = \frac{e^{2K_1}}{\rho\sigma_i^2}\left(\int_{M_{1i}} \widehat{\mu}_{ji}dj - \lambda m_{1i}rp_i\right),$$

where M_{1i} is the set of sophisticated investors learning about asset i and $m_{1i} \in [0, 1]$ denotes the fraction of sophisticated investors learning about asset i. Using the conditional distribution of signals, $\int_{M_{1i}} \widehat{\mu}_{ji}dj = \lambda m_{1i}\left[\overline{z} + (1 - e^{-2K_1})\varepsilon_i\right]$, and $\int_{M_{1i}} q_{ji}dj = \lambda m_{1i}\left[\frac{\overline{z} - rp_i}{\rho\sigma_i^2} + \left(e^{2K_1} - 1\right)\frac{z_i - rp_i}{\rho\sigma_i^2}\right]$.

Total holdings of this asset by both informed *and* uninformed sophisticated investors are:

$$Q_{1i} = \int_{M_{1i}} q_{ji}dj + \lambda(1 - m_{1i})\frac{(\overline{z} - rp_i)}{\rho\sigma_i^2} = \lambda m_{1i}\left[\frac{\overline{z} - rp_i}{\rho\sigma_i^2} + \left(e^{2K_1} - 1\right)\frac{z_i - rp_i}{\rho\sigma_i^2}\right] + \lambda(1 - m_{1i})\frac{\overline{z} - rp_i}{\rho\sigma_i^2}$$

$$= \lambda \left(\frac{\overline{z}-rp_i}{\rho\sigma_i^2}\right) + \lambda m_{1i}\left(e^{2K_1}-1\right)\left(\frac{z_i-rp_i}{\rho\sigma_i^2}\right).$$

Per capita holdings for the sophisticated investors are

$$q_{1i} = \frac{Q_{1i}}{\lambda} = \left(\frac{\overline{z}-rp_i}{\rho\sigma_i^2}\right) + m_{1i}\left(e^{2K_1}-1\right)\left(\frac{z_i-rp_i}{\rho\sigma_i^2}\right).$$

Analogous expressions hold for the unsophisticated investors, with m_{2i} denoting the fraction of unsophisticated investors learning about asset i :

$$Q_{2i} = (1-\lambda)\left(\frac{\overline{z}-rp_i}{\rho\sigma_i^2}\right) + (1-\lambda)m_{2i}\left(e^{2K_2}-1\right)\left(\frac{z_i-rp_i}{\rho\sigma_i^2}\right),$$

$$q_{2i} = \frac{Q_{2i}}{1-\lambda} = \left(\frac{\overline{z}-rp_i}{\rho\sigma_i^2}\right) + m_{2i}\left(e^{2K_2}-1\right)\left(\frac{z_i-rp_i}{\rho\sigma_i^2}\right).$$

Finally, total sophisticated and unsophisticated demand for asset i is

$$Q_{1i} + Q_{2i} = \left(\frac{\overline{z}-rp_i}{\rho\sigma_i^2}\right) + \Phi_i\left(\frac{z_i-rp_i}{\rho\sigma_i^2}\right),$$

where we define $\Phi_i \equiv \lambda m_{1i}\left(e^{2K_1}-1\right) + (1-\lambda)m_{2i}\left(e^{2K_2}-1\right)$.

□

Proof of Lemma 2 (Equilibrium Prices). The market clearing condition for each asset in state (z_i, x_i) is

$$\int_{M_{1i}}\left(\frac{s_{ji}-rp_i}{e^{-2K_1}\rho\sigma_i^2}\right)dj + \int_{M_{2i}}\left(\frac{s_{ji}-rp_i}{e^{-2K_2}\rho\sigma_i^2}\right)dj + (1-m_{1i}-m_{2i})\left(\frac{\overline{z}-rp_i}{\rho\sigma_i^2}\right) = x_i,$$

where M_{1i} denotes the set of measure $m_{1i} \in [0,\lambda]$ of sophisticated investors who choose to learn about asset i, and M_{2i} denotes the set of measure $m_{2i} \in [0, 1-\lambda]$, of unsophisticated investors who choose to learn about asset i.

Using the conditional distribution of the signals, $\int_{M_{1i}} s_{ji}dj = m_{1i}\left[\overline{z} + \left(1-e^{-2K_1}\right)\varepsilon_i\right]$ for the type-1 investors, and analogously for the type-2 investors. Then, the market clearing condition can be written as $\alpha_1\overline{z} + \alpha_2\varepsilon_i - x_i = \alpha_1 rp_i$, where

$$\alpha_1 \equiv \frac{1+m_{1i}\left(e^{2K_1}-1\right)+m_{2i}\left(e^{2K_2}-1\right)}{\rho\sigma_i^2} \text{ and } \alpha_2 \equiv \frac{m_{1i}\left(e^{2K_1}-1\right)+m_{2i}\left(e^{2K_2}-1\right)}{\rho\sigma_i^2}.$$

We obtain identification of the coefficients in $p_i = a_i + b_i\varepsilon_i - c_i\nu_i$ as

$a_i = \frac{1}{r}\left[\overline{z} - \frac{\overline{x}}{\alpha_1}\right]$, $b_i = \frac{\alpha_2}{r\alpha_1}$, and $c_i = \frac{1}{r\alpha_1}$.

Let $\Phi_i \equiv m_{1i}\left(e^{2K_1}-1\right) + m_{2i}\left(e^{2K_2}-1\right)$ be a measure of the information capacity allocated to learning about asset i in equilibrium. Further substitution yields

$$a_i = \frac{1}{r}\left(\overline{z} - \frac{\rho\sigma_i^2\overline{x}}{1+\Phi_i}\right), \quad b_i = \frac{1}{r}\left(\frac{\Phi_i}{1+\Phi_i}\right), \quad c_i = \frac{1}{r}\left(\frac{\rho\sigma_i^2}{1+\Phi_i}\right). \qquad \square$$

Proof of Lemma 3 (Equilibrium Learning). Substituting a_i, b_i, and c_i in $G_i \equiv (1-rb_i)^2 + \frac{r^2c_i^2\sigma_x^2}{\sigma_i^2} + \frac{(\overline{z}-ra_i)^2}{\sigma_i^2}$ and defining $\xi_i \equiv \sigma_i^2\left(\sigma_x^2+\overline{x}^2\right)$ gives $G_i = \frac{1+\rho^2\xi_i}{(1+\Phi_i)^2}$.

By Lemma 1, each investor learns about a single asset among the assets with the highest gain. WLOG, assets are ordered such that $\sigma_i > \sigma_{i+1}$, for all $i \in \{1,...,n-1\}$. First suppose that all investors learn about the same asset. Since G_i is increasing in σ_i, this asset is asset 1. All investors learn about asset 1 as long as $\phi \leq \phi_1 \equiv \sqrt{\frac{1+\rho^2\xi_1}{1+\rho^2\xi_2}} - 1$. At this threshold, some investors switch and learn about the second asset.

For $\phi > \phi_1$, equilibrium gains must be equated among all assets with positive learning mass. Otherwise, investors have an incentive to switch to learning about the asset with the higher gain. Moreover, the gains of all assets with zero learning mass must be strictly lower. Otherwise, an investor would once again have the incentive to deviate and learn about one of these assets.

To derive expressions for the mass of investors learning about each asset, we assume that the participation of sophisticated and unsophisticated investors in learning about a particular asset is proportional to their mass in the population: $m_{1i} = \lambda m_i$ and $m_{2i} = (1-\lambda)\,m_i$, where m_i is the total mass of investors learning about asset i. The necessary and sufficient conditions for determining $\{m_i\}_{i=1}^n$ are $\sum_{i=1}^k m_i = 1$; $\frac{1+\phi m_i}{1+\phi m_1} = c_{i1}$, for any $i \in \{2,...,k\}$, where $c_{i1} \equiv \sqrt{\frac{1+\rho^2\xi_i}{1+\rho^2\xi_1}} \leq 1$, with equality *iff* $i=1$; and $m_i = 0$ for any $i \in \{k+1,...,n\}$. Recursively, $m_i = c_{i1}m_1 - \frac{1}{\phi}(1-c_{i1})$, $\quad \forall i \in \{2,...,k\}$. Using $\sum_{i=1}^k m_i = 1$, and defining $C_k \equiv \sum_{i=1}^k c_{i1}$, we obtain the solution for m_1 given by $m_1 = \frac{1}{C_k} + \frac{1}{\phi}\left(\frac{k}{C_k}-1\right)$. Using this expression, we obtain the solution for all m_i, $i \in \{1,...,k\}$, $m_i = \frac{c_{i1}}{C_k} + \frac{1}{\phi}\left(\frac{kc_{i1}}{C_k}-1\right)$. $\square$

1.2 Analytic Results

Proof of Proposition 1. Results follow from equations that define ownership. $\square$

Proof of Proposition 2. **(i)**Follows from the definition of capital income per capita. **(ii)** Since for all $i \in \{1, ..., k\}$, the gains G_i are equated in equilibrium, then $E[\pi_{1i} - \pi_{2i}]$ is increasing in m_i, which in turn is increasing in σ_i^2. $\square$

Proof of Proposition 3. **(i)** The increase in dispersion keeps ϕ unchanged. Therefore, the masses m_i are unchanged. With both ϕ and m_i unchanged, prices are unchanged. **(ii)** The result follows from equation (10): masses and prices do not change, and dispersion, $\left(e^{2K_1} - e^{2K_2}\right)$ increases. **(iii)** Relative capital income is

$$\frac{\pi_{1i}}{\pi_{2i}} = \frac{(\overline{z}_i - rp_i)(z_i - rp_i) + \left(e^{2K_1} - 1\right) m_i (z_i - rp_i)^2}{(\overline{z}_i - rp_i)(z_i - rp_i) + \left(e^{2K_2} - 1\right) m_i (z_i - rp_i)^2} > 1.$$

Since prices are unchanged, $(\overline{z}_i - rp_i)(z_i - rp_i)$ and $m_i (z_i - rp_i)^2$ are unchanged. Since $K_1' > K_1$ and $K_2' < K_2$, the second term in π_{1i} increases and the second term in π_{2i} decreases. $\square$

Proof of Proposition 4. **(i)** Using equilibrium prices, $\overline{p}_i = \frac{1}{r}\left(\overline{z} - \frac{\rho\sigma_i^2\bar{x}}{1+\phi m_i}\right)$. The quantity ϕm_i is increasing in ϕ. Hence, for $i \in \{1, ..., k\}$, $\overline{p}_i$ is increasing in ϕ. The result for equilibrium expected excess returns $E[z_i - r\overline{p}_i]$ follows.

(ii) Since $\lambda E[q_{1i}] + (1-\lambda) E[q_{2i}] = \bar{x}$, it is sufficient to show that for $i \in \{1, ..., k'\}$, $E[q_{1i}]$ increases in response to symmetric capacity growth. Let $K \equiv K_1$, and $K_2 = \delta K$, with $\delta \in (0, 1)$. Since

$E[q_{1i}] = \frac{1+m_i\left(e^{2K}-1\right)}{(1+\phi m_i)}\bar{x}$, then $\frac{dE[q_{1i}]}{dK} = \frac{\bar{x}}{(1+\phi m_i)^2}\left[\frac{d\left[m_i\left(e^{2K}-1\right)\right]}{dK}(1+\phi m_i) - \frac{d(\phi m_i)}{d\phi}\frac{d\phi}{dK}m_i\left(e^{2K}-1\right)\right]$.

Hence $sign\left(\frac{dE[q_{1i}]}{dK}\right) = sign\left(\frac{d\left[m_i\left(e^{2K}-1\right)\right]}{dK} - \frac{d(\phi m_i)}{d\phi}\frac{d\phi}{dK}\frac{m_i\left(e^{2K}-1\right)}{1+\phi m_i}\right)$.

The quantity $\frac{d\left[m_i\left(e^{2K}-1\right)\right]}{dK} > 2e^{2K}\frac{d(\phi m_i)}{d\phi} > 0$. Hence,

$sign\left(\frac{dE[q_{1i}]}{dK}\right) = sign\left(2e^{2K} - \frac{d\phi}{dK}\frac{m_i\left(e^{2K}-1\right)}{1+\phi m_i}\right)$

$$= sign \left(2e^{2K} - \frac{2m_i\left[\lambda e^{2K}+(1-\lambda)\delta e^{2K\delta}\right]\left(e^{2K}-1\right)}{1+m_i\left[\lambda(e^{2K}-1)+(1-\lambda)\left(e^{2K\delta}-1\right)\right]} \right)$$

$$= sign \left(e^{2K} - \left(e^{2K}-1\right) \frac{m_i\left[\lambda e^{2K}+(1-\lambda)\delta e^{2K\delta}\right]}{1+m_i\left[\lambda e^{2K}+(1-\lambda)e^{2K\delta}\right]-m_i} \right)$$

$$\overset{(1)}{=} sign \left(e^{2K} - \left(e^{2K}-1\right) \left[\frac{m_i\left[\lambda e^{2K}+(1-\lambda)e^{2K\delta}\right]}{1+m_i\left[\lambda e^{2K}+(1-\lambda)e^{2K\delta}\right]-m_i} \right] \right)$$

$$\overset{(2)}{=} sign \left(e^{2K} - \left(e^{2K}-1\right) \right) > 0$$

where (1) follows from $\delta \in (0,1)$, and (2) follows from the fact that the term in square brackets is less than 1.

(iii) Let the per capita capital income be decomposed into a component C_i that is common across investor groups, and a component that is group-specific:
$\pi_{1i} = c_i + \frac{1}{\rho\sigma_i^2} m_i \left(e^{2K}-1\right)(z_i - rp_i)^2$, where $c_i \equiv \frac{1}{\rho\sigma_i^2}\left(\overline{z} - rp_i\right)(z_i - rp_i)$, with expected value C_i. Then $E\left[\pi_{1i}\right] = C_i + \frac{1}{\rho\sigma_i^2} m_i \left(e^{2K}-1\right) E\left[(z_i - rp_i)^2\right] = C_i + \frac{1}{\rho} m_i \left(e^{2K}-1\right) G_i$, where G_i is the gain from learning about asset i, equated across all $i \in \{1, ..., k\}$.
We then obtain that $\frac{E[\pi_{1i}]}{E[\pi_{2i}]} = \frac{C_i + \frac{1}{\rho} m_i \left(e^{2K}-1\right) G_i}{C_i + \frac{1}{\rho} m_i \left(e^{2K\delta}-1\right) G_i}$.
In response to an increase in K, C_i and G_i decrease, but they affect both sophisticated and unsophisticated profits in the same way. The quantity $m_i\left(e^{2K}-1\right)$ increases by more than $m_i\left(e^{2K\delta}-1\right)$ in response to a change in K. Hence overall, $\frac{E[\pi_{1i}]}{E[\pi_{2i}]}$ increases.

□

1.3 Additional Results

1.3.1 Asymmetric Capacity Growth

How much would investor-specific capacity growth amplify inequality? In bringing the model to the data, we have linked investor sophistication to initial wealth. We now investigate the impact of this link by supposing that the growth in capacity for each investor type is proportional to that group's own returns, rather than to the market average return. Since the sophisticated investors earn higher returns, their capacity growth is also higher. This results in further dispersion in capacities. As shown in Section 3, larger dispersion amplifies inequality, but it does not affect asset prices or the average market return. In the parameterized economy, such a feedback from returns to capacity growth generates a 49% growth in inequality, nearly 30% more than the benchmark, as shown in Table 1. In this asymmetric specification, we keep the average capacity growth rate at 4.9% annually, as in the benchmark.

This exercise considers a reduced-form feedback loop, while maintaining the relationship between initial wealth and initial capacity as exogenous. Below we study how such a relation could arise endogenously, and what it implies for capacity differences between types.

1.3.2 Endogenous Capacity Choice

We provide a numerical example of an endogenous capacity choice outcome in a model in which wealth heterogeneity matters for endogenous capacity choice. In particular, we assume that investors have identical CRRA preferences with IES coefficient γ, and differ in terms of their beginning of period wealth. Then, for each investor j, the absolute risk aversion coefficient is a function of wealth W_j, given by

$$A(W_j) = \gamma / W_j.$$

Locally, we map this into absolute risk aversion differences in a mean-variance optimization model by setting the coefficient ρ_j for investor j equal to $A(W_j)$. These differences in absolute risk aversion in the model imply differences in the size of the risky portfolio, and hence different gains from investing wealth in purchases of information capacity.

In particular, for a given cost of capacity given by the function $f(K)$, each investor type is going to choose the amount of capacity to maximize the ex-ante expectation of utility:

$$\frac{1}{2\rho_j} \sum_{i=1}^{n} \frac{\sigma_i^2}{\widehat{\sigma}_{ij}^2} G_i - f(K_j),$$

where, in equilibrium, G_i is a function of the distribution of individual capacity choices of investors, but not of individual capacity choices, and $\widehat{\sigma}_{ij}^2 = \sigma_i^2 e^{-2K_j}$ if the investor learns about asset i.

The gain from increasing wealth is given by the benefit of increasing the precision of information for the asset that the investor is learning about. Since all actively traded assets have the same gain in equilibrium, we can express the marginal benefit of increasing capacity in terms of the gain of the highest volatility asset (asset 1), $\frac{1}{2\rho_j}e^{2K_j}G_1$, and then the optimization problem for capacity choice can be expressed as

$$\max_K \left\{ \frac{1}{2\rho_j} e^{2K} G_1 - f(K) \right\}. \tag{1}$$

Assumption 1 below ensures an interior solution to (1) exists.

Assumption 1. The following statements hold:
(i) For all j, $\frac{G_1}{\rho_j} - f'(0) > 0$, where G_1 is evaluated at $K_j = 0$ for all j,
(ii) There exists $\underline{K} > 0$, such that for all j and for all $K > \underline{K}$, $2\frac{G_1}{\rho_j}e^{2K} - f''(K) < 0$,
(iii) There exists $\bar{K} > 0$ such that for all j and for all $K > \bar{K}$, $\frac{G_1}{\rho_j}e^{2K} - f'(K) < 0$.

Numerical example Assume that the cost function is of the form: $f(K) = e^{aK}$. Under Assumption 1, the optimal choice of K for agent j is implicitly defined by:

$$\frac{G_1(\{\bar{K}_j\})}{\rho_j} = ae^{(a-2)K},$$

where we make the dependence of G_1 on the distribution of capacities explicit. Clearly, for any $a > 2$, the higher wealth investors (implying lower ρ_j) will choose higher capacity levels. However, because of the dependence of G on equilibrium capacity choices, to quantify the differences we need to solve the equilibrium fixed point of the model.

Figure 1 presents the ratio of capacities as a function of the cost parameter of capacity, a, for different values of the absolute risk aversion coefficient of the wealthy ρ_1 (which maps into different common relative risk aversion coefficients γ). The inequality in capacity exhibits a U-shape. First, if the cost of capacity is small, then the equilibrium inequality in capacity grows without bound, as the wealthier accumulate infinite capacity (faster than the less wealthy). For higher values of the cost of capacity, inequality exhibits a growing trend as the cost increases, very quickly approaching values in excess of 38, our benchmark value. It should be noted that even for the high values of the cost parameter, the overall cost relative to gain, $f(K_j)/\frac{1}{2\rho_j}e^{2K_j}G_1$, is relatively small, less than 1% for the wealthy and less than 6% for the less wealthy.

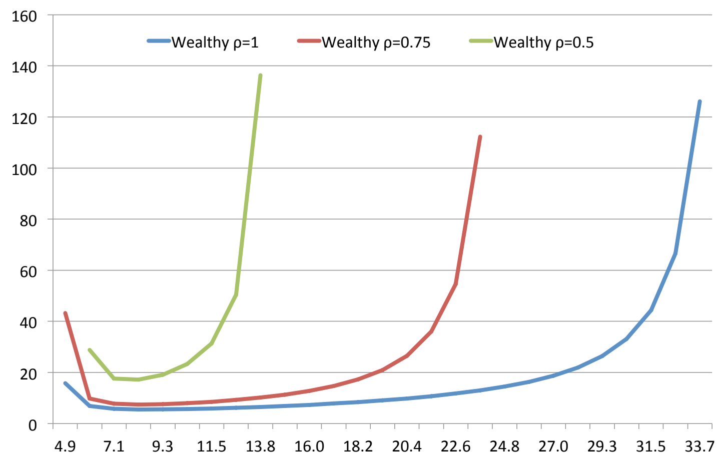


Figure 1: Inequality in information capacity (K_1/K_2) as a function of a and absolute risk aversion coefficient of the wealthy.

Intuitively, if investors endogenously choose different portfolio sizes, then their net benefit from investing in information increases with portfolio size, which further increases dispersion in capacity choice and hence portfolios.

1.3.3 CRRA Utility Specification

We also solve the main investment problem of maximizing the expected utility of wealth, where the utility function is CRRA with respect to end of period wealth:

$$\max E \frac{W^{1-\rho}}{1-\rho} \tag{2}$$

where $\rho \neq 1$. Generally, for our specification of the payoff process, i.e. $z \sim \mathcal{N}(\bar{z}, \sigma_i^2)$, wealth next period is

$$W_{t+1} = r(W_t - \sum_i p_i q_i) + \sum_i q_i z_i$$

which has a Normal distribution if z_i's are Normal. In order to analytically express the expectation in (2), we start by expressing wealth as $W' = We^{\log\{[r(1-\sum p\frac{q}{W})+\sum \frac{q}{W}z]\}}$, and then use an approximation of the log return.

Approximation To approximate $\log\{[r(1-\sum p\frac{q}{W})+\sum \frac{q}{W}z]\}$, define

$$f(z-rp) \equiv \log[r + \frac{1}{W}\sum pq\frac{z-rp}{p}].$$

In the above equation, the term z is the only unknown stochastic term. Its Taylor approximation is

$$f(z-rp) = f(\bar{z}-rp) + f'(\bar{z}-rp)(z-\bar{z}) + \frac{1}{2}f''(\bar{z}-rp)(z-\bar{z})^2 + o(z-rp)$$

where in the above,

$$f' = \frac{1}{r+\frac{1}{W}\sum q(\bar{z}-rp)}\frac{q}{W},$$

$$f'' = -\frac{1}{(r+\frac{1}{W}\sum q(\bar{z}-rp))^2}\frac{q^2}{W^2},$$

$$f''' = 2\frac{1}{(r+\frac{1}{W}\sum q(\bar{z}-rp))^3}\frac{q^3}{W^3}.$$

With these formulas in hand, the approximation is

$$f(z-rp) = \log[r + \frac{1}{W}\sum q(\bar{z}-rp)] + \frac{1}{r+\frac{1}{W}\sum q(\bar{z}-rp)}\frac{q}{W}(z-\bar{z})$$

$$-\frac{1}{2}\frac{1}{(r+\frac{1}{W}\sum q(\bar{z}-rp))^2}\frac{q^2}{W^2}(z-\bar{z})^2$$

Denote

$$r + \frac{1}{W}\sum q(\bar{z}-rp) \equiv R(q)$$

Then we can write

$$f(z-rp) = \log[R(q)] + \frac{1}{R(q)}\frac{q}{W}(z-\bar{z}) - \frac{1}{2}\frac{1}{R(q)^2}\frac{q^2}{W^2}(z-\bar{z})^2,$$

and

$$(e^{\log(f(z-rp))})^{1-\rho} = e^{(1-\rho)(\log[R(q)] + \frac{1}{R(q)}\frac{q}{W}(z-\bar{z}) - \frac{1}{2}\frac{1}{(R(q))^2}\frac{q^2}{W^2}(z-\bar{z})^2)}$$

$$= (R(q))^{1-\rho} e^{(1-\rho)\frac{1}{R(q)}\frac{q}{W}(z-\bar{z}) - \frac{1}{2}(1-\rho)\frac{1}{(R(q))^2}\frac{q^2}{W^2}(z-\bar{z})^2}$$

We are interested in the object $e^{(1-\rho)\frac{1}{R(q)}\frac{q}{W}(z-\bar{z}) - \frac{1}{2}(1-\rho)\frac{1}{(R(q))^2}\frac{q^2}{W^2}(z-\bar{z})^2}$ from the above expression. First, we approximate the term $(z-\bar{z})^2$ by its expected volatility, $\sigma_{\delta i}^2$, to get

$$e^{(1-\rho)\frac{1}{R(q)}\frac{q}{W}(z-\bar{z}) - \frac{1}{2}(1-\rho)\frac{1}{(R(q))^2}\frac{q^2}{W^2}\sigma_{\delta i}^2}$$

As an approximation point, we pick $\bar{z}$, which gives a constant $R(q)$, and then

$$\log EW^{1-\rho} = const. \times \log Ee^{(1-\rho)\frac{1}{R(q)}\frac{q}{W}(z-\bar{z}) - \frac{1}{2}(1-\rho)\frac{1}{(R(q))^2}\frac{q^2}{W^2}\sigma_{\delta i}^2} \tag{3}$$

where the variable in the exponent is Normal, with mean (ignoring constants) $\sum q_i(\hat{\mu}_i - \bar{z}_i)$ and variance equal to $\sum q_i^2 \sigma_{\delta i}^2$. Then,

$$\log EW^{1-\rho} = const. \times (1-\rho)\left\{\frac{1}{R}\sum \frac{q}{W}(\hat{\mu}_i - \bar{z}_i) + (1-\rho)\frac{1}{W^2R^2}\frac{1}{2}\sum q_i^2\sigma_{\delta i}^2\right.$$

$$\left. -\frac{1}{2}\frac{1}{W^2R^2}\sum q_i^2\sigma_{\delta i}^2\right\}$$

which gives

$$\log EW^{1-\rho} = const. \times (1-\rho)\left\{\frac{1}{R}\sum \frac{q}{W}(\hat{\mu}_i - \bar{z}_i) - \rho\frac{1}{W^2R^2}\frac{1}{2}\sum q_i^2\sigma_{\delta i}^2\right\}$$

Interior minimum (which maximizes $EW^{1-\rho}/(1-\rho)$) is

$$q_i = \frac{1}{\rho}\frac{\hat{\mu}_i - rp}{\sigma_{\delta i}^2}(Wr).$$

Plugging in gives:

$$U = \frac{1}{1-\rho}W^{1-\rho}r^{1-\rho}e^{\frac{1-\rho}{\rho}\frac{1}{2}\sum\frac{(\hat{\mu}_i - rp)^2}{\sigma_{\delta i}^2}}$$

where $\hat{\mu}_i$ and $\sigma_{\delta i}$ are the expected mean and standard deviation of the payoff process z, given the investor's prior, private signal, and the price signal.

We compute the expectation E(U) as in **?**. Some new notation is needed for that. First, denote the excess return as

$$R_i \equiv \hat{\mu}_i - rp_i$$

with mean $\hat{R}_i$. Denote the period zero volatility of $R_i - \hat{R}_i$ as $\hat{V}_i$ (which is just the volatility of R_i). Then, we can write (in a matrix form):

$$U = \frac{1}{1-\rho} W^{1-\rho} r^{1-\rho} e^{\frac{1-\rho}{\rho}\frac{1}{2}[(R-\hat{R})\Sigma_\delta^{-1}(R-\hat{R})+2\hat{R}\Sigma_\delta^{-1}(R-\hat{R})+\hat{R}\Sigma_\delta^{-1}\hat{R}]}$$

Which gives

$$EU = \frac{1}{1-\rho} W^{1-\rho} r^{1-\rho} |I - 2\hat{V}\frac{1-\rho}{2\rho}\Sigma_\delta^{-1}|^{-1/2} \times$$

$$\exp(\frac{(1-\rho)^2}{2\rho^2}\hat{R}\Sigma_\delta^{-1}(I - 2\hat{V}\frac{1-\rho}{2\rho}\Sigma_\delta^{-1})^{-1}\hat{V}\hat{R}\Sigma_\delta^{-1} + \frac{1-\rho}{2\rho}\hat{R}\Sigma_\delta^{-1}\hat{R})$$

and

$$EU = \frac{1}{1-\rho} W^{1-\rho} r^{1-\rho} (\Pi_i(1 - \hat{V}_i \frac{1-\rho}{\rho}\sigma_{\delta i}^{-1}))^{-1/2} \times \exp\left(\frac{1-\rho}{2\rho}\sum \frac{\hat{R}_i^2}{\sigma_{\delta i}}\left[(1+\frac{\hat{V}_i}{\sigma_{\delta i}}\frac{\rho-1}{\rho})^{-1}\right]\right).$$

Logging the negative of that and simplifying gives

$$-\log(-EU) = const. + \frac{1}{2}\sum_i \log(1 + \frac{\hat{V}_i}{\sigma_{\delta i}}\frac{\rho-1}{\rho}) + \frac{\rho-1}{2\rho}\sum_i \frac{\hat{R}_i^2}{\sigma_{\delta i} + \hat{V}_i\frac{\rho-1}{\rho}}$$

This objective function is strictly decreasing in $\sigma_{\delta i}$ and convex, which means that agents are going to invest all capacity into learning about one asset. For that asset, $\sigma_{\delta i} = e^{-2K}\sigma_{yi}$, and $\sigma_{\delta i} = \sigma_{yi}$ otherwise.

1.3.4 Expansion of Asset Space

The last several decades have been marked by changes in idiosyncratic risk in the U.S. economy. To explore the role that such changes might play in the dynamics of income inequality, we consider an expansion of the assets available for investment. For illustrative purposes, we introduce new assets at the high end of the volatility spectrum, with each new asset being 1% more volatile than the previous highest-volatility asset. The emergence of these new assets actually *reduces* the growth in capital income inequality, as shown in the Table 1. High-volatility assets make the information processing more difficult, making effective capacity lower. In response, the ownership shares of sophisticated investors grow less rapidly and the price impact is reduced, resulting in higher market returns. This general equilibrium effect amplifies the direct effect of lower effective capacity, leading to more moderate capital income inequality growth. Because the volatility of the asset market is growing in this exercise, excess returns are higher for a given rate of capacity growth. We consider a reparameterization of the model that increases the rate of aggregate capacity growth to 7.4%, to match the decline in the market return seen in the data. In that case, the growth in capital income inequality is 85%—much higher than in the benchmark model—as sophisticated investors take advantage of the now more volatile asset set. The sophisticated investors hold 34% more of high-volatility assets and 30% more of low-volatility assets, relative to their population weights, at the end of the simulation. This result illustrates the strong link between structural change in the economy and inequality in capital income.

Table 1: Aggregate Capacity Growth Outcomes

	Baseline	New assets		Asym. growth
Capacity growth	4.9	4.9	7.4	5.1
Average market return	7	10	7	5.6
Capital income ineq. growth	38	28	85	54
Sophis end own share of top	1.21	1.17	1.34	1.24
Sophis end own share of bottom	1.14	1.08	1.30	1.17

1.3.5 Skill versus Risk

How much of the growth in inequality comes from differences in exposure to risk versus differences in skill? Fagereng et al. (2016b) document that risk taking is only partially responsible for the difference in returns among Norwegian households, with approximately half of the return difference being attributed to unobservable heterogeneity. Our model is one in which both risk-taking differences and pure compensation for skill generate return heterogeneity. Sophisticated investors are more exposed to risk because they hold a larger share of risky assets (compensation for risk); and they have informational advantage (compensation for skill). To shed more light on the relative importance of these two effects, we decompose the returns of each investor type by computing the unconditional expectation of the return on the portfolio held by investor type $j \in \{S, U\}$:

$$R_j = E\sum_i \omega_{jit}(r_{it} - r) = \sum_i Cov(\omega_{jit}, r_{it}) + \sum_i E\omega_{jit}E[r_{it} - r], \tag{4}$$

where $r_{it} = z_{it}/p_{it}$ is the time t return on asset i and ω_{jit} is the portfolio weight of asset i for investor j at time t, defined as $\omega_{jit} = q_{jit}p_{it}/\sum_l q_{jlt}p_{lt}$. The first term of the decomposition captures the covariance conditional on investor j information set, that is, the investor's reaction to information flow via portfolio weight adjustment (*skill effect*); the second term captures the *average effect*, unrelated to active trading.

Quantitatively, the skill effect accounts for the majority of the return differential in the model. To show that, we compute the counterfactual return of sophisticated investors if their skill matched that of unsophisticated (plus noise) investors, but their average holdings stayed the same

$$\hat{R}_I = \sum_i Cov(\omega_{Rit}, r_{it}) + \sum_i E\omega_{Iit}E[r_{it} - r]. \tag{5}$$

Such a portfolio would have generated an annualized return of 10.3%, which implies that the compensation for skill accounts for more than 75% of the return differential between the sophisticated and unsophisticated investors.

2 Appendix: Additional Data Discussion and Analysis

Our model parametrization is based on the data from the Survey of Consumer Finances (SCF) from 1989 to 2013. For all computed statistics, we weigh all observations by the weights provided by the SCF (variable `42001`). Consistent with previous studies we drop farm owners.

2.1 Data Constructs

Participation Our measure of participation in financial markets includes individuals who satisfy at least one of the following criteria: (i) have a brokerage account (coded in variable `3923`), (ii) report a positive amount of stock holdings (variable `3915`), (iii) report holding non money market funds (coded as a positive balance in at least one of the variables: `3822`, `3824`, `3826`, `3828`, and `3830`; and also `7787` starting in survey year 2004), (iv) report positive holdings of bonds (coded as the sum of: `3906`, `3908`, `3910`, and additionally `7633`, `7634` starting in survey year 1992), (v) report dividends from their stock holdings (variable `5710`), (vi) report holding money funds (coded in variables: `3507`, `3511`, `3515`, `3519`, `3523`, `3527`). As a robustness check, we also consider a measure of broad market participation that includes the above six, plus the condition that a household has equity in a retirement account. Specifically, we consider the criterion that (vii) a household reports that either the head or spouse or other family members have money in retirement accounts invested in equity. For survey years 1989 and 1992 it is coded in variable `3631` with values of 2 (stocks, mutual funds), 4 (combination of stocks, CDs and money market accounts, and bonds), 5 (combination of stocks and bonds), 6 (combination of CDs and money market accounts, and stocks). For survey years 1995, 1998, and 2001 it is coded in variable `3631` with values of 2, 4, 5, 6, or 16 (brokerage account/cash management account). For surveys starting in 2004, the coding shifts to variables `6555`, `6563`, or `6571` (head, spouse, other family members). For the 2004, 2007, and 2013 surveys, this means values 1 (all in stocks), 3 (split), or 5 (hedge fund) for at least one of the variables. For survey year 2010, this means answering 1, 3, 5, or 30 (mutual fund). Adding the category of -7 (other) to the above list does not change the results.

Capital Income To construct a measure of capital income, we sum up income from four sources: (i) dividend income (5710), (ii) income from non-taxable investments such as municipal bonds (5706), (iii) net gains or losses from mutual funds, sale of stocks, bonds, or real estate (5712), and (iv) other interest income (5708).

Wealth Measures Total wealth is a sum of financial and non-financial wealth as per **?**. Financial wealth is a sum of: (1) holdings in non money funds (sum of balance in variables: 3822, 3824, 3826, 3828, 3830, and also 7787 starting in survey year 2004), (2) bond holdings balance (the sum of: 3906, 3908, 3910, and also 7633, 7634 starting in survey year 1992), (3) balance of directly held stocks (variable 3915), (4) cash value of life insurance (4006), (5) other financial assets (future royalties, money owed to households, etc. in variable 4018), (6) balances in individual retirement accounts of all family members (variables 6551-6554, 6559-6562, 6567-6570, 6756, 6757, 6758), (7) value of certificates of deposit (3721), (8) cash value of annuities, trusts, or managed accounts (6577, 6587), (9) value of savings bonds (3902), (10) value of liquid assets (checking accounts 3506, 3510, 3514, 3518, 3522, 3526, 3529, cash or call money accounts 3930, savings and money market accounts 3730, 3736, 3742, 3748, 3754, 3760).

Non-financial wealth is a sum of: (1) value of vehicles, including motor homes, RVs, motorcycles, boats, and airplanes less the amount still owed on the financing loans for these vehicles (8166+8167+8168+8188-2218-2318-2418-7169+2506+2606-2519-2619+2623-2625), (2) value of business in which a household has either active or nonactive interest (value of active business is calculated as net equity if business was sold today plus loans from the household to the business minus loans from the business to the household, plus value of personal assets used as collateral for business loans; value of non-active business is the market value; the formula used (for the 2004 SCF) is 3129+3229+3329+3335+8452+8453+3408+3412+3416+3420 +3424+3428+3124+3224+3324-(3126+3226+3326) plus 3121+3221+3321 (variables have different numbers pre-1995; some variables are not reported in 2010 and 2013 anymore), (3) value of houses and mobile homes/sites owned (604+614+623+716), (4) value of other real estate owned: vacation homes (2002) and owned share of other property (1706*1705+1806*1805+1906*1905 divided by 10000), (5) the value of other non-residential real estate net of mortgages and other loans taken out for investment in real estate (2012-2016), (6) other non-financial assets, such as artwork, precious metals, antiques, oil and gas leases, futures contracts, future proceeds from a lawsuit or estate that is being settled, royalties, or something else (4022+4026+4030).

Wage Income and Total Income For labor income and total income, we use the SCF responses to questions 5702 (income from wages and salaries) and 5729 (income from all sources). The difference between the two, apart from capital income, consists of social security and other pension income, income from professional practice, business or limited partnerships, income from net rent, royalties, trusts and investment in business, unemployment benefits, child support, alimony and income from welfare assistance programs.

2.2 Participation

In Figure 1, we present the time series of our two measures of participation. The series *Participation* follows our benchmark definition above, while *Participation + Retirement* is a broader measure that also includes individuals who participate in equity through retirement accounts.

Our participation measure changes from 32% in 1989 to a high of 40% in 2001, and down to 28% in 2013. When we additionally include participation through retirement accounts, the dynamics are very similar, except that the levels get shifted upwards. The participation level is around 35% in 1989, peaks at 44% in 2001, and goes down to 37% in 2013.

Even though both measures of participation exhibit considerable variation over time (although without any particular trend), as we point out in the paper, financial wealth inequality in the SCF data set is entirely concentrated within our participating group. Figure 2 in the paper, reproduced in Figure 3 below, presents financial wealth inequality between

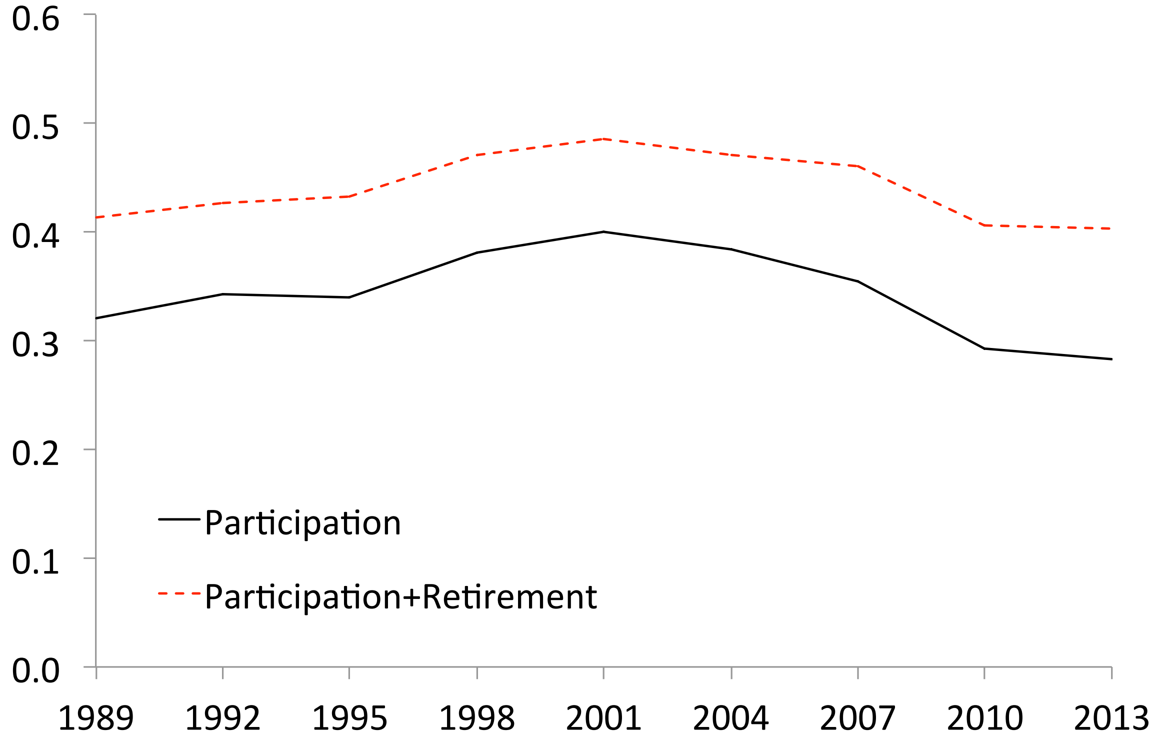


Figure 2: Financial markets participation in the SCF.

(i) top decile versus the rest of our participating group ('Sophisticated/Unsophisticated'), (ii) bottom decile of participants and non-participants, and (iii) bottom decile of participants and non-participants. Financial wealth inequality between the bottom participants and non-participants, exhibits no trend and the ratios are stable around 1. Additionally, also in Figure 2 in the paper, we show that all of the growth in financial wealth inequality in our participating group can be accounted for by retained capital income. These two points suggest that the participating group is the relevant subsample to study capital income inequality.

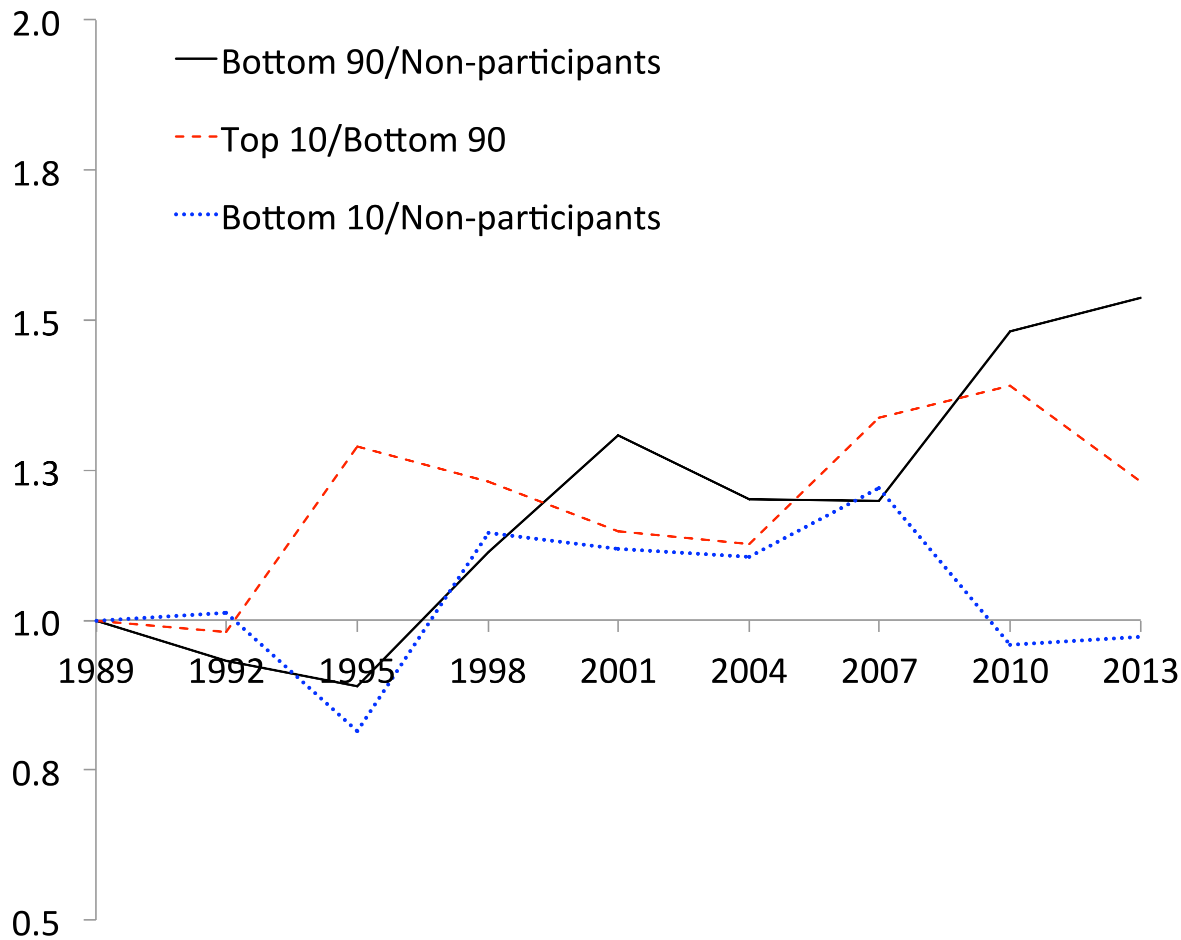


Figure 3: Extensive and intensive margins in capital income inequality.

2.3 Capital Income

Inequality Our measure of inequality is the mean income in the top decile of the wealth distribution relative to the mean income in the rest (*of participants*). Figure 4 presents the evolution of capital income inequality in the SCF. Figure 5 presents the evolution of capital income inequality using the benchmark definition of participation as well as Participation+Retirement.

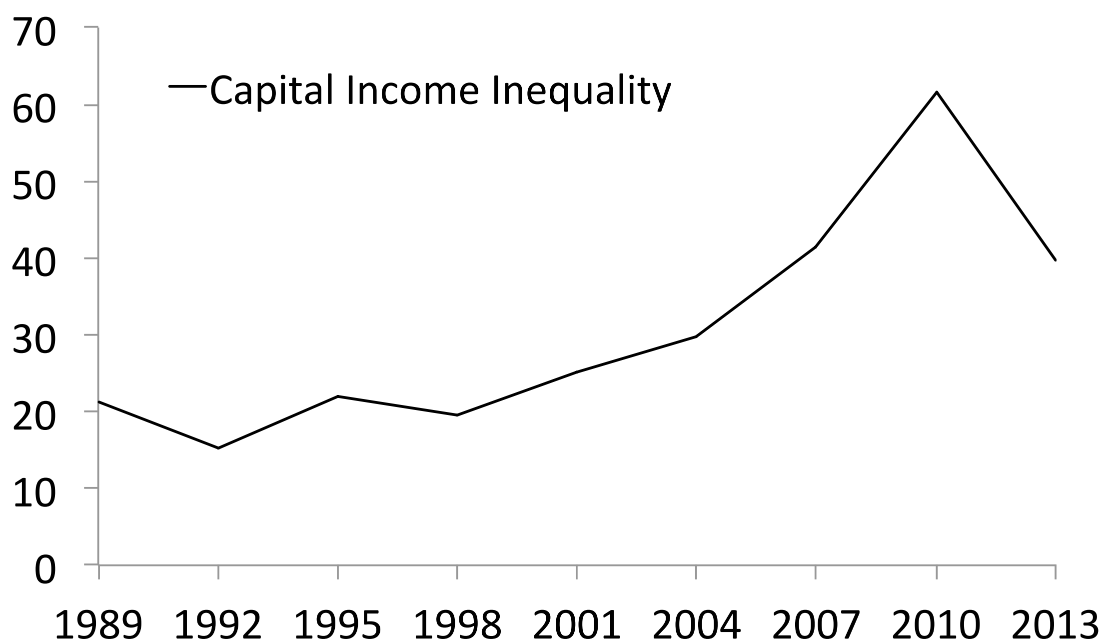


Figure 4: Capital income inequality.

Passive Investment Policies We also study whether capital income differences are an outcome of time-varying *market returns* combined with *passive* buy-and-hold household strategies. It is possible that some households (the wealthy) hold a larger share of their wealth in stocks, which gives them higher returns by the mere fact that stocks outperform bonds. In Figure 6 we plot, for each year, the past 15-year cumulative return on holding the aggregate index of the U.S. stock market.[1] We contrast this return with that of a household exclusively holding bonds (with a gross return of 1).

[1]The patterns we document are essentially the same for other choices of the horizon: 5, 10, or 20 years.

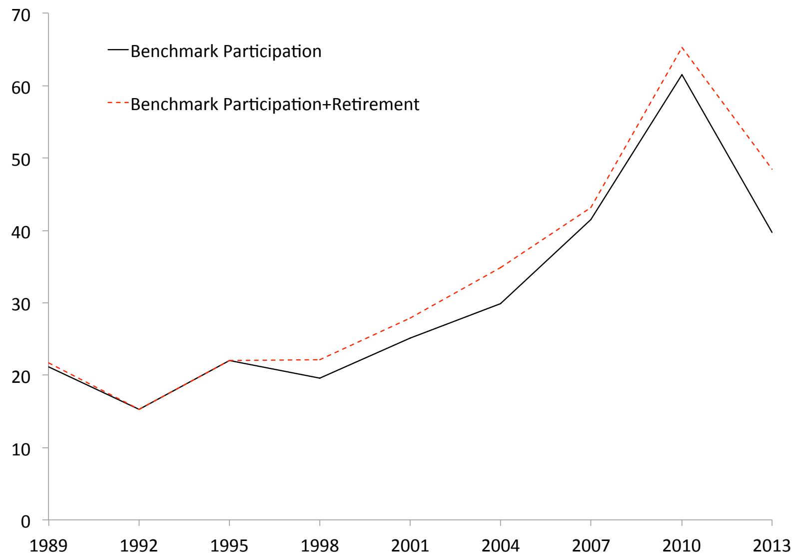


Figure 5: Capital income inequality for different measures of participation.

The cumulative return on the passive strategy exhibits a declining trend, which implies that if investors used the passive strategy and the only difference was how much money they hold in the stock market versus bonds, then we should observe a declining trend in capital income inequality, as the gross return on the market converges to the gross return on bonds. This exercise highlights the importance of active decisions of when to enter and exit the stock market.

2.4 Survey of Consumer Finances: Descriptive Statistics

To complete the characterization of the participating and non-participating groups in the SCF, Table 2 presents summary statistics for the 1989 and 2013 surveys. As expected, participants in financial markets tend to be wealthier, older and more educated. Within the participating group, the top 10% of participants also have higher financial wealth, are older, and more educated. However, Panel I of the table shows that the growth in financial wealth inequality is concentrated almost exclusively within the participating group, consistent with the trends in Figure 3. First, in the cross-section, the financial wealth of the bottom 50% of participants is only twice that of the non-participants; conversely, in 1989 the top 10% of participants has financial wealth that is 38 times larger than that of the bottom 50%.

Figure 6: Cumulative market return on a 15-year passive investment in the U.S. stock market.

Second, between 1989 and 2013, financial wealth inequality within the participants group grew by 67% (top 10% versus bottom 50%), while inequality across groups (bottom 50% versus Non-participants) grew by mere 12%. Panels II through IV of Table 2 summarize the inequality in capital, labor, and total income for participants and non-participants. The same pattern that emerged with respect to financial wealth inequality also applies to labor and total income inequality: both the level and the growth of inequality have been concentrated within the group of participants.

Panels V through VIII of the table explore potential drivers of the growth in inequality between the top 10% and the bottom 90% or 50% of participants. First, top participants hold a much smaller fraction of their financial wealth in liquid assets (Panel V). In turn, bottom participants start out with a higher share (28% or 33% versus 21%) and also grow the fraction of financial wealth held in liquid assets significantly (from 28% and 33% in 1989 to 37% and 46% in 2013). This type of portfolio composition shift towards lower risk liquid assets for the bottom participants is consistent with our information-based mechanism. Third, top participants also have higher educational attainment and are much more likely to have brokerage accounts (Panels VI and VII), consistent with their having a higher degree of financial sophistication. The data, however, also show a significant increase in access to brokerage accounts for the bottom participants (from 25% and 16% in 1989 to 46% and 35%

Table 2: Investor Characteristics in the SCF

	1989	2013
I. Financial Wealth		
Top 10%/Bottom 90% of Participants	13	16
Bottom 90%/Non-participants	5.8	8.8
Bottom 50%/Non-participants	2	2.2
II. Capital Income		
Top 10%/Bottom 90% of Participants	21	39.7
Bottom 90%/Non-participants	-	-
III. Wages and Salaries Income		
Top 10%/Bottom 90% of Participants	2.4	3.9
Bottom 90%/Non-participants	1.8	2.
Bottom 50%/Non-participants	1.3	1.4
IV. Total Income		
Top 10%/Bottom 90% of Participants	5.6	7.2
Bottom 90%/Non-participants	1.9	2
Bottom 50%/Non-participants	1.25	1.27
V. Liquid Assets/Financial Wealth		
Top 10% of Participants	21%	19%
Bottom 90% of Participants	28%	37%
Bottom 50% of Participants	33%	46%
Non-participants	52%	75%
VI. Has brokerage account		
Top 10% of Participants	64%	83%
Bottom 90% of Participants	25%	46%
Bottom 50% of Participants	16%	36%
VII. % with college		
Top 10% of Participants	67%	87%
Bottom 90% of Participants	40%	56%
Bottom 50% of Participants	31%	46%
Non-participants	15%	23%
VIII. Age (years)		
Top 10% of Participants	57	60
Bottom 90% of Participants	51	54
Bottom 50% of Participants	49	51
Non-participants	46	50

Source: SCF. Capital income/Financial wealth is the ratio of average capital income to the average financial wealth in each group. Percent with college is the fraction of individuals with 16 or more years of schooling. See the Online Appendix for complete definitions.

in 2013). This fact, along with evidence that transaction costs on brokerage accounts have been trending down (?), suggests that the costs of accessing and transacting in financial markets are an unlikely explanation for the observed rise in capital income inequality. If anything, the improved access to financial markets should generate lower inequality, in the absence of informational heterogeneity. Finally, while top participants are on average older, there are no time-series dynamics to the age difference that could explain the observed capital income dynamics (Panel VIII).

2.5 Mutual Funds and Delegation

Barriers to High-Return Institutional Funds We compare returns from different types of mutual funds, using data from Morningstar, which contains information for two types of funds: those with a minimum investment of $100,000 (institutional funds) and those without such restrictions (retail funds). Our fund data span the period 1989 through 2012.

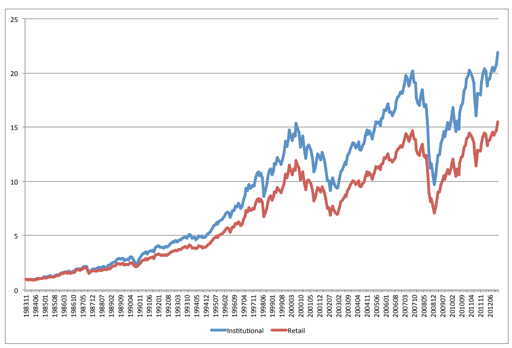


Figure 7: Cumulative investment returns in equity mutual funds by investor type.

Figure 7 plots the cumulative return series for institutional versus retail mutual funds. To construct the figure, we compute the value of one dollar invested in each fund type in January of 1989 and assume that the monthly after-fee return is subsequently reinvested until December 2012. A cumulated value of the dollar in 1989 grows to $22 for institutional funds and to $16 for retail funds. This difference amounts to about 3% return difference per year between the two types of funds. Since the institutional funds have a minimum investment threshold, less sophisticated, less wealthy investors do not have access to the higher returns earned by institutional funds, even for "plain vanilla" assets like equities.

Dispersion in the Quality of Asset Management Companies We document a large heterogeneity in mutual fund returns in the data depending on the investment size (typically related to wealth). Additionally, below we show that the average fund does not outperform the passive benchmark and that the performance of a typical mutual fund is not persistent over time. Taken together, these findings suggest that selecting a mutual fund in any particular period is an informationally intensive task, similar to trading individual stocks.[2]

Average Mutual Fund Does Not Outperform Passive Benchmark. We construct a sample of risk-adjusted after-fee fund returns by regressing monthly excess fund returns, net of the risk-free rate, on four risk factors: market, size, value, and momentum as in Carhart (1997). The abnormal return from this regression is our definition of a risk-adjusted return. We present in Figure 8 a histogram of monthly returns pooled across all funds and all months in our sample. The mean and median value of the distribution are not statistically different from zero.

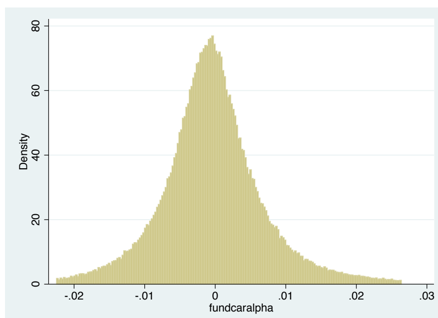


Figure 8: Distribution of equity funds' returns.

Mutual Fund Performance Is Not Persistent Over Time. While an average fund does not beat a passive benchmark, we observe a large cross-sectional dispersion in returns with both small and large values of alpha. It is thus possible that investors could focus their attention only on funds with positive returns thus beating the market portfolio. The issue with such approach is whether funds with positive returns tend to outperform the benchmark on a consistent basis. If not, the strategy of focusing on current winners may not be profitable.

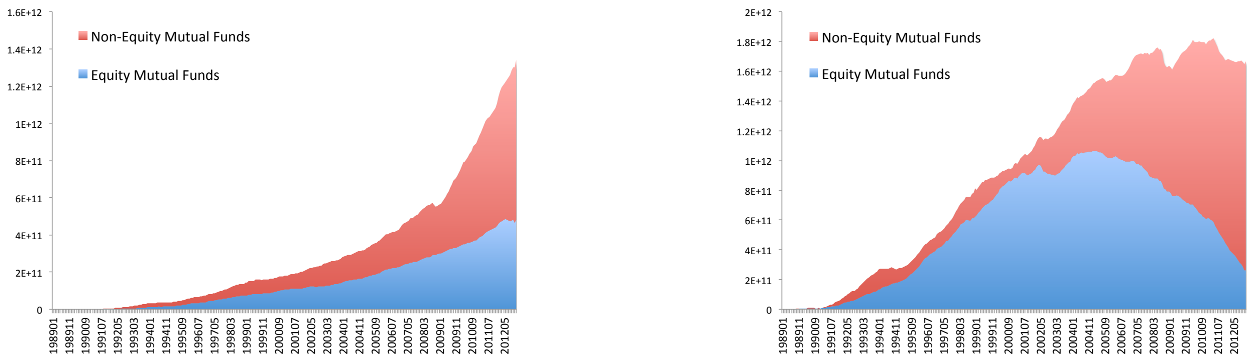
[2]We are not the first ones to point out these regularities. Extant literature in finance, such as **?** or **?** finds that while the average abnormal gross returns of mutual funds are positive, the distribution of returns is highly dispersed and the returns are not predictable.

To test for such predictability, each month we sort funds into five equal-sized portfolios according to their current risk-adjusted returns and test whether the ranking of funds into such portfolios is preserved one month and one year into the future. We show the result using a transition matrix of being in a particular quintile portfolio conditional on starting in a given portfolio at time t. Each of the 25 cells of the transition matrix illustrates the probability of being in quintile $j = 1 - 5$ at time $t + k$ conditional on being in quintile $i = 1 - 5$ at time t. We set k to be equal to 1 and to 12 months. The results are in Table 3.

Table 3: Transition Probabilities of Fund Performance

Performance quintiles					
	at $t + k$				
at t	1	2	3	4	5
	k=1 month				
1	77.8	16.6	3.4	1.3	0.7
2	16.5	56.2	20.7	5.2	1.4
3	3.6	20.5	52.1	20.3	3.4
4	1.2	5.4	20.3	56.6	16.4
5	0.8	1.5	3.6	16.6	77.4
	k=12 months				
1	29.3	19.6	16.6	16.6	17.8
2	20.0	22.2	21.6	20.3	15.9
3	16.6	22.0	23.7	21.9	15.7
4	16.0	21.0	21.9	22.1	19.1
5	18.5	16.7	17.0	19.8	27.9

We observe that fund performance is not very persistent over time. For example, a fund that starts in the top-performing quintile at time t has a 77% chance of ending up in the same quintile one month later. The same probability for one-year ahead transition drops to 28%. Similar patterns emerge for other quintiles in the matrix. We conclude that an uninformed household would face a difficult task to invest in a successful fund by simply following past winners.



(a) Institutional (b) Retail

Figure 9: Cumulative Flows to Mutual Funds: Institutional vs. Retail

2.6 Expansion of Ownership

As aggregate capacity grows, sophisticated investors expand their ownership of risky assets by order of volatility: starting from the highest volatility assets and then moving down.

To test this prediction, we consider flows into mutual funds by investor type. Given that equity funds are generally more risky than non-equity funds one would expect unsophisticated investors be less likely to invest in the equity funds, especially if aggregate information capacity grows.

We use data on flows into equity and non-equity mutual funds from Morningstar by sophisticated (institutional) and unsophisticated (retail) investors. As shown in Figure 9, the cumulative flows from sophisticated investors into equity and non-equity funds increase steadily over the entire sample period. In contrast, the flows from unsophisticated investors display a markedly different pattern. The flows into equity funds grow until 2000 but subsequently decrease at a significant rate to drop by a factor of 3 by 2012. Moreover, this decrease coincides with a significant increase in cumulative flows to non-equity funds. Notably, the increase in equity fund flows by unsophisticated investors observed in the early sample period is consistent with the steady decrease in holdings of individual equity in the U.S. data. To the extent that direct equity holdings are more risky than diversified equity portfolios, such as mutual funds, this implies that unsophisticated investors have been systematically reallocating their wealth from riskier to safer asset classes.

Overall, these findings qualitatively support our model's predictions: Sophisticated households have a large exposure to risky assets and subsequently add exposure to less risky assets, and as unsophisticated households face greater information disadvantage they increasingly move their money into safer assets.