

Investor Sophistication and Capital Income Inequality*

Marcin Kacperczyk Jaromir Nosal Luminita Stevens
Imperial College London & CEPR Boston College University of Maryland

August 14, 2017

Abstract

Capital income inequality is large and growing fast, accounting for a significant portion of total income inequality. We study its determinants in a general equilibrium portfolio choice model with endogenous information acquisition and heterogeneity across household sophistication and asset riskiness. The main mechanism works through endogenous household participation in assets with different risk. The model implies capital income inequality that increases with aggregate information technology. Quantitatively, the model generates growth of capital income inequality that accounts for a significant fraction of the growth in inequality in the U.S. data.

*We thank Boragan Aruoba, Bruno Biais, Laurent Calvet, John Campbell, Bruce Carlin, John Donaldson, Thierry Foucault, Xavier Gabaix, Mike Golosov, Gita Gopinath, Jungsuk Han, Ron Kaniel, Kai Li, Matteo Maggiori, Gustavo Manso, Alan Moreira, Stijn van Nieuwerburgh, Stavros Panageas, Alexi Savov, John Shea, Laura Veldkamp, and Venky Venkateswaran for useful suggestions and Joonkyu Choi for research assistance. Kacperczyk acknowledges research support by a Marie Curie FP7 Integration Grant within the 7th European Union Framework Programme and by European Research Council Consolidator Grant. Contact: m.kacperczyk@imperial.ac.uk, jarek.nosal@gmail.com, stevens@econ.umd.edu.

1 Introduction

The rising trends in income and wealth inequality have been among the most hotly discussed topics in academic and policy circles.¹ Among the various possible explanations, heterogeneity in the rates of return on savings has been highlighted as an important driver of the distribution of both income and wealth in the data. This factor has emerged in empirical work that has studied the entire wealth distribution, such as Fagereng, Guiso, Malacrino, and Pistaferri (2016b, 2016a), as well as in research that has focused on the very top of the wealth distribution (Benhabib, Bisin, and Zhu (2011)) and inequality due to entrepreneurship forces (Pástor and Veronesi (2016)).² However, as noted by De Nardi and Fella (2017), more work is needed to understand the determinants of such heterogeneity.

In this paper, we aim to address this gap by quantifying how much inequality can arise from a general equilibrium theory of portfolio choice in which rate of return heterogeneity arises endogenously due to investor heterogeneity, and in which returns and investment patterns are disciplined by data from financial markets. We model the effects of differences in investor skill on the evolution of capital income inequality in a model with endogenous information acquisition. We show how, given skill heterogeneity, general progress in information technology can result in increased capital income inequality, as it disproportionately benefits the investors who have access to high skill investment technologies. When disciplined by data from financial markets and the Survey of Consumer Finances, the parameterized model generates a quantitatively significant growth in capital income inequality, accounting for nearly 50% of

¹See Piketty and Saez (2003); Atkinson, Piketty, and Saez (2011). A comprehensive discussion is also offered in the 2013 Summer issue of the *Journal Economic Perspectives* and in Piketty (2014).

²See also the review by Benhabib and Bisin (2017). A notable exception in this literature is the work of Saez and Zucman (2016), who emphasize the role of differential savings rates over that of differential rates of return. While both mechanisms are plausible in principle, the capitalization method these authors use imposes homogeneity in the rates of return within asset classes, thereby assuming one mechanism over the other. See also the critique of Fagereng, Guiso, Malacrino, and Pistaferri (2016b) who show that imposing this homogeneity assumption can overstate the degree of wealth inequality when it is violated.

the data.

At the core of our model is each investor’s decision of how much to invest in assets with different risk characteristics. This decision is shaped by the investor’s capacity to process information about asset payoffs. We model the learning choice using the theory of rational inattention (Sims (2003)). In this framework, investors endowed with a fixed capacity to learn about different asset payoffs must decide how to allocate their capacity: which assets to learn about, how much information about them to process, and how much to invest.³ Our theoretical framework generalizes existing models by considering heterogeneous informed investors facing multiple heterogeneous assets.⁴

In the model, we analytically characterize a rich set of testable predictions about heterogeneity in investor portfolios that we then evaluate against the financial data. In particular, we show that investors with higher capacities to process information about risky assets hold larger portfolios on average and they additionally invest more in the riskier assets within the portfolio of risky assets. Hence, both *portfolio size* and *portfolio composition* differ across investors.⁵ These patterns are consistent with the empirical literature on portfolio composition differences between wealthy and less wealthy investors, going back to Greenwood (1983), Kessler and Wolff (1991), and Mankiw and Zeldes (1991), and shown more recently by Fagereng, Guiso, Malacrino, and Pistaferri (2016b).⁶ This, together with work that has linked trading strategy

³In the model, we endow each investor with a particular level of information processing capacity. However, this capacity should be interpreted more broadly, as a stand-in for the individual’s ability to access high quality investment advice, not limited to his or her own knowledge of or ability to invest in financial markets.

⁴Prior models using rational inattention in an asset market context, such as Van Nieuwerburgh and Veldkamp (2009, 2010), or Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) typically focus on polar cases of heterogeneity in investor capacity of the uninformed/informed type, or a single risky asset case. Our framework extends the model discussed in the appendix of Van Nieuwerburgh and Veldkamp (2010) to a heterogeneous investor case.

⁵Investors in our model trade all assets, but as is typical in this type of framework, they specialize in learning about one asset. The model, therefore, generates underdiversification due to information frictions. Underdiversification of portfolio holdings has been documented in the literature (see Vissing-Jorgensen (2004) and references therein). In the aggregate, however, learning is endogenously spread across multiple assets. For more details, see Section 2.

⁶Bach, Calvet, and Sodini (2015) also document portfolio composition heterogeneity, using

sophistication to asset prices, wealth and income levels,⁷ motivates our quantitative exercise of mapping information capacity differences into wealth deciles in the data.

The main result of the paper is the effect on income inequality of *symmetric* growth in information capacity, interpreted as a general progress in information-processing technologies. We show that such progress disproportionately benefits the initially more skilled investors and leads to growing capital income inequality. This result reflects two characteristics of learning in equilibrium. First, learning exhibits *preference for volatility*: All else equal, individuals choose to learn about more volatile assets. Second, there is *strategic substitutability* in learning: The value of learning diminishes as more individuals learn about a given asset, through a general equilibrium effect on prices. Less sophisticated individuals are more responsive to the general equilibrium price effects because their information rents are lower. As a result, symmetric growth in capacity leads to an expansion of sophisticated ownership across asset classes, starting with the most volatile and continuing to lower volatility assets. Simultaneously, unsophisticated individuals retrench from risky assets and hold safer assets. In terms of aggregates, general progress in information technology also generates lower market returns, higher market turnover, and larger and more volatile portfolios. We document that these predictions are borne out in data from CRSP and Morningstar on stocks and mutual funds over the last 25 years.

To evaluate the quantitative importance of these effects, we parameterize the model using financial data on returns and asset volatilities, and data from the Survey of Consumer Finances on asset holdings and risky rates of return differences. To establish the quantitative goal for the theory, we measure capital income inequality growth in the SCF between year 1989 and 2013.⁸ Specifically, we focus on the subset of

Swedish data, and they show that this heterogeneity is a major contributor to the financial wealth inequality that exists in their data, though they attribute less of it to skill.

⁷See Calvet, Campbell, and Sodini (2009), Chien, Cole, and Lustig (2011), and Vissing-Jorgensen (2004).

⁸The SCF provides very detailed, high quality data on the balance sheets of a representative sample of U.S. households. As shown by Saez and Zucman (2016), the SCF displays a top 10% wealth share that is very close to that obtained using administrative tax data. Where the SCF falls

households that participate in risky financial markets (roughly 34% of the population in the SCF), and split those households into the top decile and the bottom 9 deciles of their total wealth. Motivated by the discussion of investor heterogeneity above, we map the sophisticated and unsophisticated⁹ households in the model into the top decile and the bottom 9 deciles of wealth, respectively. We then simulate the model for 25 years, under a symmetric capacity growth rate equal to 5.1% annually, which is selected to match the dynamics of the overall market excess return in the model. We find that the model implies growth in capital income inequality of 42% versus 87% growth in the data, accounting for 48% of the data increase. Importantly, we find that in an analogously parameterized model with a single risky asset, growth in capital income inequality is 20%, accounting for 23% of the data. Hence, more than half of the quantitative effect in the benchmark model is due to increasing differences in portfolio composition in the presence of asset heterogeneity. We conclude our analysis with a set of additional dynamic predictions, relating to data on cross-sectional turnover and the expansion of ownership by asset type.

The crucial driver of the quantitative results are changes in information capacity, which are difficult to measure empirically, especially for different investor types. Given that, we focus on a case of aggregate capacity growth, which affects all investors' constraints in the same way. We then use the model as a measuring tool in our calibration exercise for determining the differences between capacities and the growth rate of capacity over time. Our 5.1% calibrated growth rate also falls in the range of values for the increase in the number of stocks actively analyzed by the financial industry (4%) and the number of analysts per stock (8%), which are two potential candidates for a measure of aggregate information capacity. Notably, due to the convexity of the capacity constraints in the model, aggregate capacity growth, even though it affects all constraints in a symmetric way, will translate into a larger

short is at estimating the top 0.1% wealth share, but this is less of a concern for our purposes, since we are not focusing on inequality at the very top of the distribution. The appendix presents detailed statistics for the top decile versus the bottom 90% and for participants versus non-participants.

⁹We also include noise traders in the unsophisticated group.

growth in precision of beliefs for the initially more sophisticated investors. However, our results are predominantly driven by the endogenous response of *portfolio composition* of different investor types rather than mechanical differences in signal precision. This is illustrated in the single asset case, in which we shut down the portfolio composition channel and show that less than half of the benchmark model’s capital income inequality growth is left.

Our model extends the work of Van Nieuwerburgh and Veldkamp (2009, 2010), and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) into modeling multiple heterogeneous risky assets and multiple investors with heterogeneous, but positive capacity. Those papers focus their analysis on polar cases of investors heterogeneity of the informed (positive capacity) and uninformed (zero capacity) types.

The idea that matching the inequality in outcomes observed in the data requires connecting rates of return to wealth is certainly not new. For example, Aiyagari (1994) discusses the wide disparities in portfolio compositions across the wealth distribution, focusing on the fact that rich households are much more likely to hold risky assets, such as equities and risky debt instruments. Subsequently, Krusell and Smith (1998) suggest that the data requires making rich agents have higher propensities to save, generate higher returns on savings, or both. Benhabib, Bisin, and Zhu (2011) develop this relationship theoretically. We build on this literature by explicitly linking the evolution of inequality to developments in financial markets. Our focus on skill rather than risk tolerance differences is supported by portfolio-level evidence provided by Fagereng, Guiso, Malacrino, and Pistaferri (2016a), who calculate that only a quarter of the difference in returns between wealthy and less wealthy households can be attributed to higher risk taking.

Our modeling of differences in rates of return on financial assets contributes to the large literature that has sought to generate inequality in capital income, including the work on bequests by Cagetti and De Nardi (2006), on limited stock market participation by Guvenen (2007, 2009), on heterogeneous discount factors by Krusell and

Smith (1998), on financial literacy by Lusardi, Michaud, and Mitchell (2017), and on entrepreneurial talent by Quadrini (1999). Our approach to generating this heterogeneity via differences in information contributes a novel mechanism to this literature, building on the insights of Arrow (1987). A complementary mechanism is presented by Pástor and Veronesi (2016) who develop a single-asset model with heterogeneity in skill and risk aversion. However, they study the link between redistributive taxes, entrepreneurship, and income inequality, and they do not model endogenous information acquisition. Another related paper is Peress (2004) who examines the role of wealth and decreasing absolute risk aversion in investors' information acquisition and investment in a single risky asset. However, his focus is not on capital income inequality. Moreover, we show that having heterogeneity across assets and agents is a crucial component to quantitatively capture the evolution of capital income inequality and its underlying economic mechanism. Relative to the literature focused on the tails of the income distribution, such as Gabaix, Lasry, Lions, and Moll (2016), we provide a mechanism that works on the entirety of the distribution and is not solely operational asymptotically.

The rest of the paper is organized as follows. Section 2 presents the theory. Section 3 derives analytic predictions, which we subsequently take to the data. Section 4 presents our empirical targets, for both inequality and asset returns. Section 5 presents the quantitative results and tests the mechanism using a set of dynamic predictions. Section 6 concludes. All proofs and derivations are in the appendix.

2 Theoretical Framework

This section presents a noisy rational expectations portfolio choice model in which investors are constrained in their capacity to process information about asset payoffs. The setup departs from the information choice model of Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) by introducing heterogeneity in both assets and investor

capacities. Moreover, in the equilibrium, we solve for the optimal mass of agents learning about each asset and we characterize the properties of learning in response to changes in the aggregate information capacity in the economy.

2.1 Setup

A continuum of atomless investors of mass one, indexed by j , solve a sequence of portfolio choice problems, so as to maximize mean-variance utility over wealth W_j in each period, given common risk aversion coefficient $\rho > 0$. The financial market consists of one risk-free asset, with price normalized to 1 and payoff r , and $n > 1$ risky assets, indexed by i , with prices p_i , and independent payoffs $z_i = \bar{z} + \varepsilon_i$, with $\varepsilon_i \sim \mathcal{N}(0, \sigma_i^2)$. The risk-free asset has unlimited supply, and each risky asset has fixed supply, \bar{x} . For each risky asset, non-optimizing “noise traders” trade for reasons orthogonal to prices and payoffs (e.g., liquidity, hedging, or life-cycle reasons), such that the net supply available to the (optimizing) investors is $x_i = \bar{x} + \nu_i$, with $\nu_i \sim \mathcal{N}(0, \sigma_x^2)$, independent of payoffs and across assets.¹⁰

Prior to making their portfolio decisions, investors choose to obtain information about some or all of the risky assets. Mass $\lambda \in (0, 1)$ of investors, labeled *sophisticated*, have high capacity to process information, K_1 , and mass $1 - \lambda$, labeled *unsophisticated*, have low capacity, K_2 , with $0 < K_2 < K_1 < \infty$. Information is obtained in the form of endogenously designed signals on asset payoffs subject to this capacity limit. The signal choice is modeled using entropy reduction as a measure of the amount of acquired information (see Sims (2003)).

2.2 Investor optimization

Optimization occurs in two stages. In the first stage, investors solve their information acquisition problem: they choose the distribution of their individual signals

¹⁰For simplicity, we introduce heterogeneity only in the volatility of payoffs, although the model can easily accommodate heterogeneity in supply and in mean payoffs.

in order to maximize expected utility, subject to their information capacity. In the second stage, given the signals they receive, investors update their beliefs about the payoffs and choose their portfolio holdings to maximize utility. We first describe the optimal portfolio choice in the second stage, for a given signal choice. We then solve for the ex-ante optimal signal choice.

Portfolio choice Given equilibrium prices and posterior beliefs, each investor's portfolio problem is standard. The investor solves

$$U_j = \max_{\{q_{ji}\}_{i=1}^n} E_j(W_j) - \frac{\rho}{2} V_j(W_j) \quad (1)$$

$$s.t. \quad W_j = r \left(W_{0j} - \sum_{i=1}^n q_{ji} p_i \right) + \sum_{i=1}^n q_{ji} z_i, \quad (2)$$

where E_j and V_j denote the mean and variance conditional on investor j 's information set, and W_{0j} is initial wealth. Optimal portfolio holdings are given by

$$q_{ji} = \frac{\hat{\mu}_{ji} - r p_i}{\rho \hat{\sigma}_{ji}^2}, \quad (3)$$

where $\hat{\mu}_{ji}$ and $\hat{\sigma}_{ji}^2$ are the mean and variance of investor j 's posterior beliefs about the payoff z_i .

Information acquisition choice Each investor j can choose to receive a separate signal s_{ji} on each of the asset payoffs z_i . Given the optimal portfolio choice, each investor chooses the optimal distribution of signals to maximize the ex-ante expected utility, $E_{0j}[U_j]$. The choice of the vector of signals $s_j = (s_{j1}, \dots, s_{jn})$ about the vector of payoffs $z = (z_1, \dots, z_n)$, is subject to an information capacity constraint, $I(z; s_j) \leq K_j$, where $I(z; s_j)$ denotes the Shannon (1948) mutual information, quantifying the information that the vector of signals conveys about the vector of payoffs.¹¹ The capacity

¹¹We solve the model for the case in which investors do not learn from prices. In the Online Appendix, we prove that this is in fact optimal if processing the information content of prices is also

constraint imposes a limit on the amount of uncertainty reduction that the signals can achieve. Since perfect information requires infinite capacity, each investor faces some residual uncertainty about the realized payoffs.

For tractability, we assume that the signals s_{ji} are independent across assets and investors. This assumption implies that the total quantity of information obtained by an investor can be expressed as a sum of the quantities of information obtained for each asset. The information constraint becomes $\sum_{i=1}^n I(z_i; s_{ji}) \leq K_j$, where $I(z_i; s_{ji})$ measures the information conveyed by the signal s_{ji} about the payoff of asset i .

Investors decompose each payoff into a lower-entropy signal component and a residual component that represents the information lost through this compression: $z_i = s_{ji} + \delta_{ji}$. To maintain analytical tractability, the signal s_{ji} is independent of the information that is lost δ_{ji} , for each asset and investor.¹² Since z_i is normally distributed, this assumption implies that s_{ji} and δ_{ji} are also normally distributed. By Cramer's Theorem, $s_{ji} \sim \mathcal{N}(\bar{z}, \sigma_{s_{ji}}^2)$ and $\delta_{ji} \sim \mathcal{N}(0, \sigma_{\delta_{ji}}^2)$ with $\sigma_i^2 = \sigma_{s_{ji}}^2 + \sigma_{\delta_{ji}}^2$. Hence, posterior beliefs are normally distributed random variables, independent across assets, with mean $\hat{\mu}_{ji} = s_{ji}$ and variance $\hat{\sigma}_{ji}^2 = \sigma_{\delta_{ji}}^2$. A perfectly precise signal results in no information loss, such that posterior uncertainty is zero. Conversely, a signal that consumes no information capacity discards all information about the realized payoff, returns only the mean payoff, \bar{z} , and leaves an investor's posterior uncertainty equal to her prior uncertainty.

The investor's information problem is then choosing the variance of posterior beliefs to solve

$$\max_{\{\hat{\sigma}_{ji}^2\}_{i=1}^n} \sum_{i=1}^n G_i \frac{\sigma_i^2}{\hat{\sigma}_{ji}^2} \quad s.t. \quad \prod_{i=1}^n \frac{\sigma_i^2}{\hat{\sigma}_{ji}^2} \leq e^{2K_j}, \quad (4)$$

costly. Intuitively, prices are an inferior source of information compared with the private signals, because the private signals are optimally designed by the investor to provide information specifically about payoffs, while prices are contaminated with information about the noise trader shocks, which are not payoff-relevant per se.

¹²The two simplifying assumptions in our setup are standard in the literature. In principle, the model can be solved numerically without making these assumptions, but analytics would not be feasible in those cases.

where G_i represents the equilibrium utility gain from learning about asset i .¹³ This gain is a function of equilibrium prices and hence it is common across investor types and taken as given by each investor.

Lemma 1. *The solution to the maximization problem (4) is a corner: each investor allocates her entire capacity to learning about a single asset from the set of assets with maximal utility gains. For all other assets, the investor's optimal portfolio holdings are determined by her prior beliefs.*

For each investor j learning about asset $l_j \in \arg \max_i G_i$, the posterior beliefs are normally distributed, with mean and variance given by

$$\hat{\mu}_{ji} = \begin{cases} s_{ji} & \text{if } i = l_j \\ \bar{z} & \text{if } i \neq l_j \end{cases} \quad \text{and} \quad \hat{\sigma}_{ji}^2 = \begin{cases} e^{-2K_j} \sigma_i^2 & \text{if } i = l_j \\ \sigma_i^2 & \text{if } i \neq l_j. \end{cases} \quad (5)$$

For $i = l_j$, conditional on the realized payoff z_i , the signal is normally distributed with mean $E(s_{ji}|z_i) = \bar{z} + (1 - e^{-2K_j}) \varepsilon_i$, and variance $V(s_{ji}|z_i) = (1 - e^{-2K_j}) e^{-2K_j} \sigma_i^2$.

The linear objective function and the convex constraint imply that each investor specializes, learning about a single asset. She always picks an asset with the highest gain G_i and hence all assets that are learned about in equilibrium will have the same gains. Which assets these are is determined in equilibrium.

2.3 Equilibrium

Equilibrium prices Given the solution to each investor's portfolio and information problem, equilibrium prices are linear combinations of the shocks.

Lemma 2. *The price of asset i is given by $p_i = a_i + b_i \varepsilon_i - c_i \nu_i$, with*

$$a_i = \frac{1}{r} \left[\bar{z} - \frac{\rho \sigma_i^2 \bar{x}}{(1 + \Phi_i)} \right], \quad b_i = \frac{\Phi_i}{r(1 + \Phi_i)}, \quad c_i = \frac{\rho \sigma_i^2}{r(1 + \Phi_i)}, \quad (6)$$

¹³The investor's objective omits terms that do not affect the optimization. See the Appendix for detailed derivations of this and all subsequent results.

where $\Phi_i \equiv m_{1i}(e^{2K_1} - 1) + m_{2i}(e^{2K_2} - 1)$ measures the information capacity allocated to learning about asset i in equilibrium, $m_{1i} \in [0, \lambda]$ is the mass of sophisticated investors who choose to learn about asset i , and $m_{2i} \in [0, 1 - \lambda]$ is the mass of unsophisticated investors who choose to learn about asset i , with $\sum_{i=1}^n m_{1i} = \lambda$ and $\sum_{i=1}^n m_{2i} = 1 - \lambda$.

The price of an asset reflects the asset's payoff and the supply shocks, with relative importance determined by the mass of investors learning about the asset. If there is no information capacity in the economy ($K_1 = K_2 = 0$), or for assets that are not learned about ($m_{1i} = m_{2i} = 0$), the price only reflects the supply shock ν_i . As the capacity allocated to an asset increases, the asset's price co-moves more strongly with the underlying payoff (c_i decreases and b_i increases, though at a decreasing rate). In the limit, as $K_j \rightarrow \infty$, the price approaches the discounted realized payoff, z_i/r , and the supply shock becomes irrelevant for price determination.

Equilibrium learning Using equilibrium prices, we now determine the assets that are learned about and the mass of investors learning about each asset. Without loss of generality, let assets be ordered such that $\sigma_i > \sigma_{i+1}$ for all $i \in \{1, \dots, n-1\}$. Let $\xi_i \equiv \sigma_i^2(\sigma_x^2 + \bar{x}^2)$ summarize the properties of asset i . The gain from learning about asset i is given by

$$G_i = \frac{1 + \rho^2 \xi_i}{(1 + \Phi_i)^2}. \quad (7)$$

Lemma 3. *The allocation of information capacity across assets, $\{\Phi_i\}_{i=1}^n$, is uniquely pinned down by equating the gains from learning among all assets that are learned about, and by ensuring that all assets not learned about have strictly lower gains:*

$$G_i = \max_{h \in \{1, \dots, n\}} G_h, \quad \forall i \in \{1, \dots, k\}, \quad (8)$$

$$G_i < \max_{h \in \{1, \dots, n\}} G_h, \quad \forall i \in \{k+1, \dots, n\}, \quad (9)$$

where k denotes the endogenous number of assets with strictly positive learning mass.

Let m_i denote the total mass of investors learning about asset i and let $c_{i1} \equiv \sqrt{\frac{1+\rho^2\xi_i}{1+\rho^2\xi_1}} \leq 1$ denote the exogenous value of learning about asset i relative to asset 1 (excluding strategic substitutability effects). In a symmetric equilibrium in which $m_{1i} = \lambda m_i$ and $m_{2i} = (1 - \lambda) m_i$, the masses $\{m_i\}_{i=1}^n$ are given by

$$m_i = \frac{c_{i1}}{C_k} + \frac{1}{\phi} \left(\frac{kc_{i1}}{C_k} - 1 \right), \quad \forall i \in \{1, \dots, k\}, \quad (10)$$

$$m_i = 0, \quad \forall i \in \{k+1, \dots, n\}, \quad (11)$$

where $C_k \equiv \sum_{i=1}^k c_{i1}$, and $\phi \equiv \lambda (e^{2K_1} - 1) + (1 - \lambda) (e^{2K_2} - 1)$ is a measure of the total capacity for processing information available in the economy, with $\Phi_i = \phi m_i$.

The model uniquely pins down the total capacity allocated to each asset, Φ_i , but it does not separately pin down m_{1i} and m_{2i} . Since the asset-specific gain from learning is the same for both types of investors, we assume that the participation of sophisticated and unsophisticated investors in learning about each asset is proportional to their mass in the population. In turn, this implies a unique set of masses $\{m_i\}_{i=1}^n$.

Learning in the cross section We turn now to characterizing how investors learn about the different assets in equilibrium, which in turn determines how much they invest in different assets. First, learning in the model exhibits preference for volatility (high σ_i^2) and strategic substitutability (low m_i). Furthermore, the value of learning about an asset also falls with the aggregate amount of information in the market (ϕ), since higher capacity overall increases the co-movement between prices and payoffs, thereby reducing expected excess returns:

$$\frac{\partial G_i}{\partial \sigma_i^2} > 0, \quad \frac{\partial G_i}{\partial m_i} < 0, \quad \frac{\partial G_i}{\partial \phi} < 0.$$

These properties imply that with a sufficiently low information capacity, all investors learn about the same asset, namely the most volatile one: for $\phi \in (0, \phi_1]$,

$m_1 = 1$ and $m_i = 0$ for all $i > 1$, where

$$\phi_1 \equiv \sqrt{\frac{1 + \rho^2 \xi_1}{1 + \rho^2 \xi_2}} - 1. \quad (12)$$

This threshold endogenizes single-asset learning as an optimal outcome for low enough information capacity relative to asset dispersion. For higher capacity levels, strategic substitutability in learning pushes some investors to learn about less volatile assets. For sufficiently high information capacity (or alternatively, for low enough dispersion in assets volatilities), all assets are actively traded, thus endogenizing the assumption employed in models with exogenous signals. Here asset heterogeneity is critical: even if capacity is high enough so that multiple assets are learned about, not all assets are learned about with the same intensity, so that holdings differ across assets, as we show below.

Learning over time We now study how learning changes in response to changes in aggregate capacity in the economy. It is useful to define the thresholds for learning as follows:

Definition 1. Let the aggregate market capacity ϕ_k be such that for any $\phi \leq \phi_k$, at most the first k assets are actively traded (learned about) in equilibrium, while for $\phi > \phi_k$, at least the first $k + 1$ assets are actively traded in equilibrium.

Lemma 3 implies that the threshold values of aggregate information capacity are monotonic: $0 < \phi_1 < \phi_2 < \dots < \phi_{n-1}$.

Lemma 4. Let $\phi \in (\phi_{k-1}, \phi_k]$ such that $k > 1$ assets are actively traded. Consider an increase in ϕ such that $k' \geq k$ is the new number of actively traded assets.

(i) There exists a threshold asset $\bar{i} < k'$, such that m_i is strictly decreasing in ϕ for all $i \in \{1, \dots, \bar{i} - 1\}$ and strictly increasing in ϕ for all $i \in \{\bar{i} + 1, \dots, k'\}$.

(ii) The quantity (ϕm_i) is increasing in ϕ for all assets $i \in \{1, \dots, k'\}$.

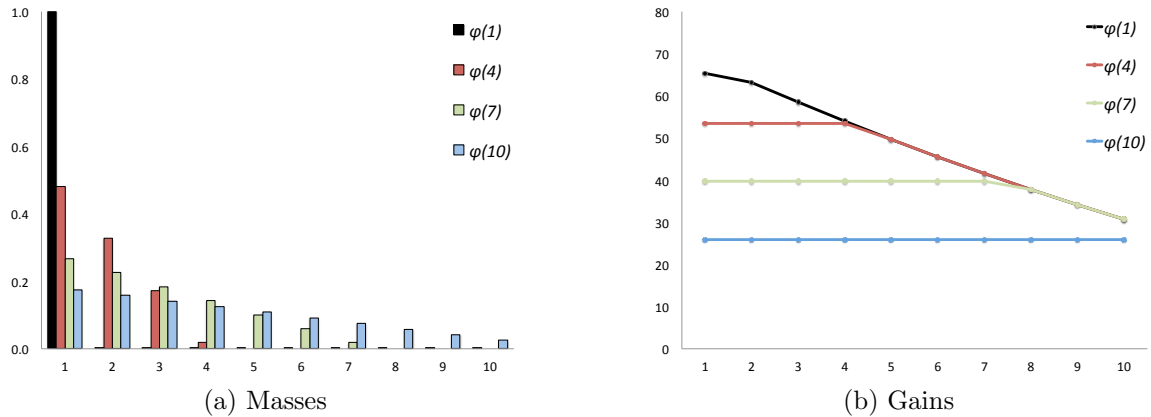


Figure 1: The evolution of masses and gains from learning as aggregate capacity is increased. $\phi(k)$ indicates the level of aggregate capacity for which k assets are learned about in equilibrium. On the x-axis, assets are ordered from most (1) to least (10) volatile.

(iii) For an increase in ϕ generated by a symmetric growth, $K'_j = (1 + \gamma) K_j$, with $\gamma \in (0, 1)$, the quantity $m_i(e^{2K_j} - 1)$, $j \in \{1, 2\}$, is increasing in K_j at an increasing rate, for $i \in \{\bar{i} + 1, \dots, k'\}$. For $i \in \{1, \dots, \bar{i}\}$, $m_i(e^{2K_1} - 1)$ grows while $m_i(e^{2K_2} - 1)$ grows by less, or even falls if capacity dispersion is large enough.

Lemma 4 shows the diversification in learning effect. First, as the amount of aggregate capacity increases, some investors shift to learning about less volatile assets, and the mass of investors learning about the most volatile assets decreases. The threshold \bar{i} determines this turning point in the distribution of assets. Figure 1 shows this effect numerically in panel (a), as the aggregate information capacity increases from ϕ_1 , the level of capacity for which only a single asset is learned about, to ϕ_{10} , the level for which ten assets are learned about.¹⁴ Nevertheless, the total amount of capacity allocated to each asset (ϕm_i) strictly increases, such that all gains from learning decline and are equated at a new, lower level for all assets that are learned

¹⁴One could also let the degree of dispersion in asset payoff volatilities vary, which will imply that learning also varies, with periods with high dispersion being characterized by more concentrated learning, and periods with low dispersion characterized by more diversified learning (and hence portfolios).

about, as shown in panel (b) of the figure. Most importantly, the increase in aggregate capacity benefits the sophisticated group disproportionately more because across all actively traded assets, this group now allocates relatively more capacity to each asset, as a result of the convexity of the entropy function. This relative capacity increase in turn generates asymmetry in investment patterns. In Section 3, we use this result to derive analytic predictions on the patterns of investment in response to changes in capacity, which we then confirm in the data.

3 Model Predictions

In this section, we present analytic results implied by our information friction.

Heterogeneous Capacity Our first set of analytic results identify the channels through which heterogeneity in information capacity drives capital income inequality in the cross-section. We show that heterogeneity in information implies differences in portfolio sizes, different composition of the risky portfolio across investors, and also implies that investors are able to adjust their holdings in response to payoff shocks more effectively if their capacity is higher. These predictions are consistent with prior empirical evidence on household portfolio heterogeneity, such as Fagereng, Guiso, Malacrino, and Pistaferri (2016a) or Bach, Calvet, and Sodini (2015).

Let q_{1i} and q_{2i} denote the average per-capita holdings of asset i for sophisticated and unsophisticated investors, respectively. They are given by

$$q_{1i} = \left(\frac{\bar{z}_i - rp_i}{\rho\sigma_i^2} \right) + m_i (e^{2K_1} - 1) \left(\frac{z_i - rp_i}{\rho\sigma_i^2} \right), \quad (13)$$

and q_{2i} defined analogously. Equation (13) shows that per-capita holdings are given by the quantity that would be held under the investors' prior beliefs plus a quantity that is increasing in the realized excess return. The weight on the realized excess return is asset and investor specific, and it is given by the amount of information

capacity allocated to this asset by this investor group. Hence, for actively traded assets, heterogeneity in capacities generates differences in ownership across investor types at the asset level:

$$q_{1i} - q_{2i} = m_i (e^{2K_1} - e^{2K_2}) \left(\frac{z_i - rp_i}{\rho\sigma_i^2} \right). \quad (14)$$

Integrating over the realizations of the state (z_i, x_i) , the expected per-capita ownership difference, as a share of the supply of each asset, is also asset specific,

$$\frac{E[q_{1i} - q_{2i}]}{\bar{x}} = (e^{2K_1} - e^{2K_2}) \frac{m_i}{1 + \phi m_i}. \quad (15)$$

Hence, the portfolio of the sophisticated investor is not simply a scaled up version of the unsophisticated portfolio. Rather, the portfolio weights *within* the class of risky assets also differ across the two investor types.

Proposition 1 (Ownership). *Let $K_1 > K_2$ and $\phi_{k-1} \leq \phi < \phi_k$, such that the first $k > 1$ assets are actively traded in equilibrium. Then, for $i \in \{1, \dots, k\}$,*

- (i) $E[q_{1i} - q_{2i}] / \bar{x} > 0$;
- (ii) $E[q_{1i} - q_{2i}] / \bar{x}$ is increasing in $E[z_i - rp_i]$ and in σ_i^2 ;
- (iii) $q_{1i} - q_{2i}$ is increasing in $z_i - rp_i$.

The average sophisticated investor (i) holds a larger portfolio of risky assets on average, (ii) tilts her portfolio towards more volatile assets with higher expected excess returns, and (iii) adjusts ownership, state by state, towards assets with higher *realized* excess returns. These results identify the channels through which capital income differs across investor types.

To see the effects of the portfolio scale and composition differences on capital income, we define the capital income of an investor of type j as $\pi_{ji} \equiv q_{ji} (z_i - rp_i)$. For actively traded assets, heterogeneity in ownership generates heterogeneity in capital

income across investor types at the asset level:

$$\pi_{1i} - \pi_{2i} = m_i (e^{2K_1} - e^{2K_2}) \frac{(z_i - rp_i)^2}{\rho\sigma_i^2}. \quad (16)$$

Integrating over the realizations of (z_i, x_i) , the expected capital income difference is

$$E[\pi_{1i} - \pi_{2i}] = \frac{1}{\rho} m_i (e^{2K_1} - e^{2K_2}) G_i, \quad (17)$$

where G_i is the gain from learning about asset i .

Proposition 2 (Capital Income). *Let $K_1 > K_2$ and $\phi_{k-1} \leq \phi < \phi_k$, such that the first $k > 1$ assets are actively traded in equilibrium. Then, for $i \in \{1, \dots, k\}$,*

- (i) $\pi_{1i} - \pi_{2i} \geq 0$, with a strict inequality in states with non-zero realized excess returns;
- (ii) $E[\pi_{1i} - \pi_{2i}]$ is increasing in asset volatility σ_i .

The average sophisticated investor realizes larger profits in states with positive excess returns, and incurs smaller losses in states with negative excess returns, because her holdings co-move more strongly with the realized state. Importantly, the biggest difference in profits, on average, comes from investment in the more volatile, higher expected excess return assets. It is these volatile assets that drive inequality because they generate the biggest gain from learning, and hence the biggest advantage from having relatively high learning capacity. Hence, holding capacity constant, the more volatile the asset market, the more unequal will be the distribution of capital income. When we quantitatively evaluate our model's predictions for inequality, we discipline the model by calibrating asset volatilities to the data.

Larger Capacity Dispersion Our second set of analytic results show that increased dispersion in capacities implies further polarization in holdings, which in turn leads to a growing capital income polarization.

Proposition 3 (Capacity Dispersion). *Let $K_1 > K_2$ and $\phi_{k-1} \leq \phi < \phi_k$, such that the first $k > 1$ assets are actively traded in equilibrium. Consider an increase in*

capacity dispersion of the form $K'_1 = K_1 + \Delta_1 > K_1$, $K'_2 = K_2 - \Delta_2 < K_2$, with Δ_1 and Δ_2 chosen such that the total information capacity ϕ remains unchanged. Then, for $i \in \{1, \dots, k\}$,

- (i) Asset prices and excess returns remain unchanged.
- (ii) The difference in ownership shares $(q_{1i} - q_{2i})/\bar{x}$ increases.
- (iii) Capital income gets more polarized as π_{1i}/π_{2i} increases state by state.

Intuitively, greater dispersion in information capacity implies that sophisticated investors receive relatively higher-quality signals about the fundamental payoffs, which enables them to respond more strongly to realized state.

However, while this increase generates higher inequality, it has no effect on financial markets, since keeping aggregate capacity unchanged implies that both the number of assets learned about and the mass of investors learning about each asset remain unchanged. Hence, the adjustment reflects a pure transfer of income from the relatively unsophisticated investors (who now have even lower capacity) to the more sophisticated investors (who now have even higher capacity) without any general equilibrium effects. To capture recent trends in financial markets, we next consider growth in aggregate capacity.

Symmetric Capacity Growth Our third and most important set of analytic results shows that in the presence of initial capacity dispersion, technological progress in the form of *symmetric growth* in information capacity, interpreted as general progress in information-processing technologies, leads to a disproportionate increase in the ownership of risky assets by the sophisticated investors, and to a growing capital income polarization, while simultaneously affecting asset prices.

Proposition 4 (Symmetric Growth). *Let $K_1 > K_2$ and $\phi_{k-1} \leq \phi < \phi_k$, such that the first $k > 1$ assets are actively traded in equilibrium. Consider an increase in ϕ generated by a symmetric growth in capacities to $K'_1 = (1 + \gamma)K_1$ and $K'_2 = (1 + \gamma)K_2$, $\gamma \in (0, 1)$. Let $k' \geq k$ denote the new equilibrium number of actively*

traded assets. Then, for $i \in \{1, \dots, k'\}$,

- (i) Average asset prices increase and average excess returns decrease.
- (ii) Average ownership share of sophisticated investors $E[q_{1i}]/\bar{x}$ increases and average ownership share of unsophisticated investors $E[q_{2i}]/\bar{x}$ decreases.
- (iii) Average capital income gets more unequal, as $E[\pi_{1i}]/E[\pi_{2i}]$ increases.

First, higher capacity to process information means that investors receive more accurate signals about the realized payoffs. Hence, their demand for assets co-moves more closely with the realized state, which implies that prices contain a larger amount of information about the fundamental shocks. As a result, the equilibrium implies lower average returns, larger and more volatile positions, and higher market turnover.

Second, a symmetric growth in capacity for both sophisticated and unsophisticated investors has two effects on portfolio holdings and capital income inequality: a partial equilibrium effect and a general equilibrium effect. Absent any equilibrium price adjustment, the average holdings of risky assets and the comovement between holdings and the realized state increase for both investor types. However, because growth in capacity benefits investors who already have relatively high capacity, the benefits accrue more for sophisticated investors. Further, in contrast to the case of increased dispersion, a symmetric change in information capacity affects equilibrium prices. As sophisticated investors increase their demand for risky assets, this drives up average prices, reducing the expected profits of unsophisticated investors, who in turn reduce their average holdings of risky securities.

Trading Volume The differential response to shocks of the two investor types also implies differences in trading intensity, which provides an additional set of testable implications. We divide the investors into 3 groups: (i) sophisticated investors who learn about asset i , with per capita average volume \bar{V}_i^{SL} ; (ii) unsophisticated investors who learn about asset i , with per capita average volume \bar{V}_i^{UL} ; and (iii) investors who do not learn about asset i , with per capita average volume \bar{V}_i^{NL} . For assets that are

not learned about, volume is denoted by \bar{V}_i^{ZL} . Hence, the total volume generated by the optimizing investors at the asset level is¹⁵

$$V_i = \begin{cases} \lambda m_i \bar{V}_i^{SL} + (1 - \lambda) m_i \bar{V}_i^{UL} + (1 - m_i) \bar{V}_i^{NL} & \text{if } i \text{ is learned about} \\ \bar{V}_i^{ZL} & \text{if } i \text{ is not learned about.} \end{cases} \quad (18)$$

We derive an analytic expression for the average per capita volume across states for each asset and investor group, given by

$$\bar{V}_i^g = \frac{1}{\sqrt{\pi}} \left(\sigma_{qi}^g + \sqrt{(\sigma_{qi}^g)^2 + (\sigma_{\mu i}^g)^2} \right), \quad (19)$$

where σ_{qi}^g is the cross-sectional standard deviation of holdings across investors in group g and $\sigma_{\mu i}^g$ is the variability of that group's mean holdings across states. Intuitively, trading volume is higher the more disagreement there is in the cross-section of investors and the more the group responds to shocks over time.

In turn, the degree of cross-sectional disagreement depends on how much capacity investors allocate to learning about that asset, with

$$(\sigma_{qi}^g)^2 = \begin{cases} \frac{e^{2K_g} - 1}{\rho^2 \sigma_i^2} & \text{if } i \text{ is learned about \& } g = \text{SL, UL} \\ 0 & \text{if } i \text{ is learned about \& } g = \text{NL} \\ 0 & \text{if } i \text{ is not learned about.} \end{cases}$$

while the degree to which investors adjust holdings over time depends on how much learning is allocated to the asset, both by the particular investor group and by the

¹⁵The average volume of the noise traders is exogenous, given by the standard deviation of the noise shock. Among optimizing investors, we assume that investors do not change groups over time. When we take the volume predictions to the data, we compute turnover, which is given by $T_i \equiv V_i/\bar{x}$.

market overall:

$$(\sigma_{\mu i}^g)^2 = \begin{cases} \left(\frac{e^{2K_g}}{1+\phi m_i}\right)^2 \sigma_x^2 + \left(\frac{e^{2K_g}-1-\phi m_i}{1+\phi m_i}\right)^2 \frac{1}{\rho^2 \sigma_i^2} & \text{if } i \text{ is learned about \& } g = \text{SL, UL} \\ \left(\frac{1}{1+\phi m_i}\right)^2 \sigma_x^2 + \left(\frac{\phi m_i}{1+\phi m_i}\right)^2 \frac{1}{\rho^2 \sigma_i^2} & \text{if } i \text{ is learned about \& } g = \text{NL} \\ \sigma_x^2 & \text{if } i \text{ is not learned about.} \end{cases}$$

These expressions enable us to derive a set of testable implications summarized below.

Proposition 5 (Volume). *Let $K_1 > K_2$ and $\phi_{k-1} \leq \phi < \phi_k$, such that the first $k > 1$ assets are actively traded in equilibrium. Then for assets that are learned about, $i \in \{1, \dots, k\}$, average volume is increasing in investor sophistication and is higher for investors who actively trade the asset: $\bar{V}_i^{SL} > \bar{V}_i^{UL} > \bar{V}_i^{NL}$.*

Hence, sophisticated investors generate more asset turnover, since having higher capacity to process information enables them to take larger and more volatile positions, relative to unsophisticated investors. Moreover, assets that are actively traded, in turn, have higher volumes compared with assets that are passively traded (based only on prior beliefs).

4 Empirical Results

Our model predicts that progress in information processing technology increases inequality, as sophisticated investors disproportionately benefit from the rising tide. In order to quantify the strength of this mechanism, we now present a set of empirical targets using both inequality and financial assets data.¹⁶

Capital Income Inequality Growth Our model yields inequality from investing in financial markets, hence we focus on the dynamics of income from risky financial

¹⁶Our evidence on capital income inequality reinforces existing results using more detailed U.S. and European data, e.g. Saez and Zucman (2016), Fagereng, Guiso, Malacrino, and Pistaferri (2016b) and Bach, Calvet, and Sodini (2015).

investment. We use the *Survey of Consumer Finances* (SCF) from 1989 to 2013. Although not as comprehensive as tax records data, the data from the Survey of Consumer Finances provide detailed information about the balance sheets of a representative sample of U.S. households.¹⁷

Financial wealth in the SCF contains holdings of risky assets (stocks, bonds, mutual funds), passive assets (life insurance, retirement accounts, royalties, annuities, trusts), and liquid assets (cash, checking and savings accounts, money market accounts). Since we seek to understand the role of financial markets in generating growth in inequality, we focus on households that participate in broadly defined risky investments. Specifically, we define as *participants* households that report holding stocks, bonds, mutual funds, receiving dividends, or having a brokerage account. On average, 34% of households participate, ranging between 32% in 1989, a 40% high in 2001, and a 28% low in 2013.¹⁸ For each survey year, we sort the sample of participants by the level of total wealth, and we calculate inequality as the ratio of average capital income of the top 10% to that of the bottom 90% of participants.

First, we document that inequality in total financial wealth has grown within the group of households who participate in financial markets, but it has remained essentially unchanged along the extensive margin (defined as the ratio of average financial wealth of the bottom 10% of participating households to that of the non-participating households). Thus the dynamics of financial wealth inequality do not appear to be driven by the extensive (participation) margin. These trends are shown in panel (a) of Figure 2.

Second, among participants, we show that the increase in inequality in financial wealth can be accounted for almost entirely by the accumulation of capital income

¹⁷See Saez and Zucman (2016) for a detailed comparison of the SCF to U.S. administrative tax data. In short, they find that the SCF is representative of trends and levels of inequality in the U.S., but understates inequality inside the top 1% of the wealth distribution.

¹⁸As a robustness check, we also consider a broader measure of participation that additionally includes all households with equity in a retirement account. This inclusion raises the participation rates to 35% in 1989, 44% in 2001, and 37% in 2013. All relevant conclusions remain unchanged.

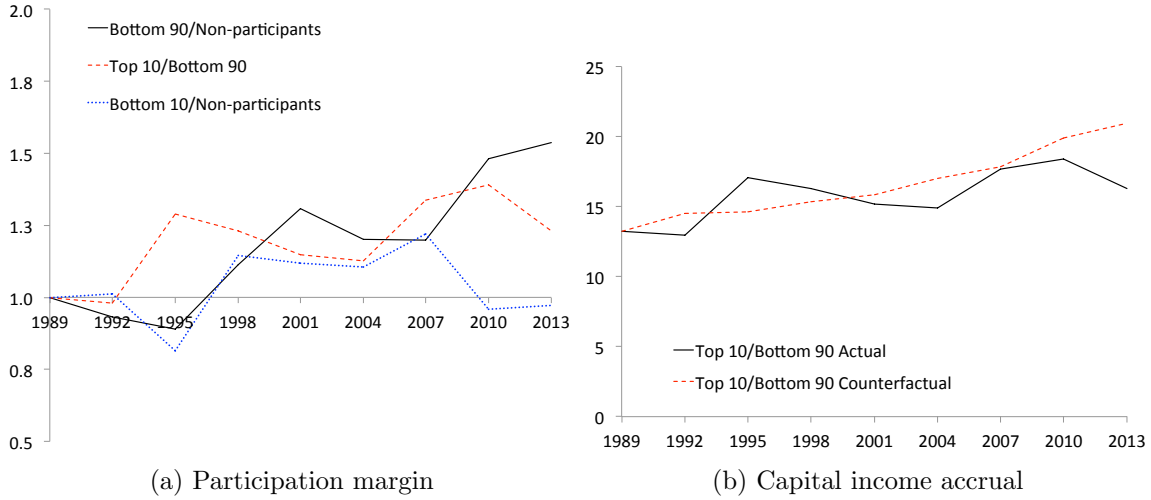


Figure 2: Financial wealth inequality in the SCF. Market participants are sorted in terms of their total wealth. (a) Inequality within the group of households who participate in financial markets, defined as the ratio of financial wealth of the top wealth decile to that of the bottom 90% of participants, versus inequality at the participating margin, measured as the ratio of financial wealth of the bottom 10% of participating households to that of the non-participating households. (b) Actual and counterfactual financial wealth inequality due to accrual of capital income only.

from the risky assets (namely, income from dividends, interest income, and realized capital gains). To see this, we consider the counterfactual financial wealth obtained from accruing capital income only.¹⁹ Panel (b) of Figure 2 suggests that past capital income realizations may be sufficient to explain the evolution of financial wealth inequality, without resorting to mechanisms that involve savings rates from other income sources.²⁰

Third, among participating households, capital income inequality is large and growing fast. Panel (a) of Figure 3 shows that in the cross-section, capital income is an order of magnitude more unequal than either labor or total income. For example,

¹⁹For example, the counterfactual financial wealth level in 1995 is equal to the actual financial wealth in 1989 plus 3 times the capital income reported in the prior survey years (in this case, 1989 and 1992).

²⁰By construction, the two wealth levels are identical in 1989, so the figure also implies that the counterfactual levels of financial wealth for each group are very close to those in the data. Still, we treat this evidence as suggestive, since our exercise imposes a panel interpretation on a repeated cross-section.

in 1989, the average capital income of the top 10% of participants was 61 times larger than that of the bottom 90% of participants. This ratio increased to 129 in 2013. By comparison, the corresponding ratio for wage income was 3.3 in 1989 and 5.6 in 2013. To compare the dynamics of inequality across income sources, we normalize the inequality of each income measure to 1 in 1989, and plot growth rates for capital, labor, and total income inequality in panel (b) of the figure. Capital income inequality nearly doubled over the sample period, outpacing the growth in labor income inequality, which increased 1.5 times.

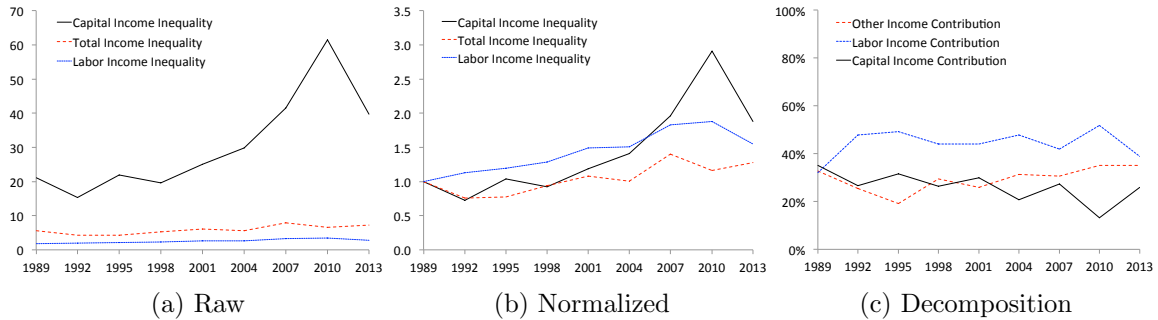


Figure 3: Income inequality growth in the SCF. Inequality is the ratio of the top 10% to the bottom 90% (in terms of total wealth) of participants in financial markets. (a) Inequality for capital income, labor income and other income in levels. (b) Same series, normalized to 1 in 1989. (c) Decomposition of total income inequality into its three components.

Since capital income is so unequal, it is an important contributor to total income inequality, even though its share in total income is not that large (14% on average). To see this, consider the decomposition of total income inequality in period t (denoted T_{10t}/T_{90t}) into shares coming from capital income (denoted by K), labor income (denoted by W), and other (residual) income (denoted by R).²¹ This pro-

²¹Labor income (income from wages and salaries) represents 56% of total income, while other income makes up the remaining 30%. Other income includes 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. In the literature on labor income inequality, business income is sometimes included in labor income. The split between labor and other income does not impact our calculations regarding the relative importance of capital income.

cess integrates two empirical drivers of inequality: the evolution of shares and the evolution of inequality within each income source.

$$\frac{T_{10t}}{T_{90t}} = \frac{K_{10t} K_{90t}}{K_{90t} T_{90t}} + \frac{W_{10t} W_{90t}}{W_{90t} T_{90t}} + \frac{R_{10t} R_{90t}}{R_{90t} T_{90t}}$$

Panel (c) of Figure 3 plots the contribution at time t of each of the components of total income to the inequality in total income. On average, 26% of the total income inequality in each year is attributable to capital income.²²

In our quantitative section, we use the series for capital income inequality growth shown in panel (b) of Figure 3 to evaluate our information-based theory of inequality. Our model predicts that investors with higher capacities hold larger portfolios of risky assets on average and moreover, *within* the portfolio of risky assets, they invest more in the riskier assets. In the data, these characteristics are correlated with initial wealth. Motivated by this correlation and by additional evidence that links wealth to sophistication (for example, Calvet, Campbell, and Sodini (2009) and Vissing-Jorgensen (2004)), we assume that an investor’s capacity is a function of wealth when mapping the model to the data.²³

Return Heterogeneity In order to quantify the return heterogeneity in the SCF, we proceed as follows. First, we calculate the holdings of risky securities for each household, comprising of holdings of stocks, bonds, and mutual funds. These are the holdings that match well with the sources of capital income in the SCF. Next, we compute the return on risky holdings as capital income divided by holdings of risky securities, and then compute the median return in the top 10% and the bottom 90% of the wealth distribution of participants. We then use these measures to capture

²²Furthermore, Benhabib, Bisin and Zhu (2011) show that it is capital income risk, not labor income risk, that is critical to generating the skewness in the wealth distribution seen in the data.

²³In the online appendix, we show that higher-wealth individuals use more sophisticated investment instruments and invest a lower proportion of their assets in money-like instruments

the heterogeneity in rates of return among the two household groups.²⁴ In particular, we compute the ratio of the return of the unsophisticated (bottom 90%) relative to the sophisticated households (top 10%), over the first half of our sample, which gives us that unsophisticated households earned 69.2% of the return of the sophisticated households. We use this heterogeneity to obtain targets for the levels of returns of each household type by requiring that the average return on risky assets be equal to the market return of 11.9% (1989-2000 average). The weights used in computing the average are the fraction of risky securities held by each type of household in the SCF (31% versus 69%). That gives us the target for sophisticated return of 13.1% and for the unsophisticated return of 9.1%. We then map these two returns into the sophisticated portfolio and the unsophisticated+noise trader portfolio in the model.²⁵

5 Quantitative Results

In this section, we parameterize the model and show that it generates the path of capital income inequality that is *quantitatively* close to the data. We also discuss additional predictions on turnover and cross-sectional asset ownership that further validate our model.

5.1 Parameterization

The complete list of parameter values and targets is presented in Table 1. The key parameters are the information capacities of the two investor types (K_1 , K_2), the fraction of sophisticated investors in the population (λ), the risk-free interest rate (r), the risk aversion parameter (ρ), the volatility of the noise shock (σ_x), and the volatilities of the payoffs, $\{\sigma_i\}_{i=1}^n$, for which we normalize the lowest volatility,

²⁴While the SCF data may not necessarily capture the levels of returns, we use them to capture the dispersion in returns among market participants.

²⁵Noise traders are also market participants from the perspective of the model, and hence would be captured by the SCF.

$\sigma_n = 1$, and assume that volatility changes linearly across assets.²⁶ Additionally, for parsimony, we restrict some parameters and normalize the natural candidates. We normalize the mean payoff to $\bar{z}_i = 10$ and asset supply to $\bar{x}_i = 5$ for all assets. We restrict the volatilities of the noise shocks, $\sigma_{xi} = \sigma_x$ for all assets, and set the number of assets to $n = 10$.²⁷

Table 1: Parameter Values

Parameter	Symbol	Value	Target (1989-2000 averages)
Mean payoff, supply	\bar{z}_i, \bar{x}_i	10, 5 for all i	Normalization
Number of assets	n	10	Normalization
Risk-free rate	r	2.5%	3-month T-bill – inflation = 2.5%
Vol. of noise shocks	σ_{xi}	0.4 for all i	Average turnover = 9.7%
Vol. of asset payoffs	σ_i	$\in [1, 1.59]$	p90/p50 of idio. return vol = 3.54
Risk aversion	ρ	1.032	Unsophisticated + noise return = 9.1%
Information capacities	K_1, K_2	0.37, 0.0037,	Sophisticated share = 69%
and investor masses	λ	0.675	Share actively traded = 50%
			Sophisticated return = 13.1%

The parameter values are chosen to jointly match key moments from the data for the first half of our sample, 1989-2000. We set the following targets: *(i)* the equity ownership share of sophisticated investors of 69%, which is the mean share of the risky assets held by the households in top 10% of the wealth distribution²⁸; *(ii)* the average return on 3-month Treasury bills minus the inflation rate, equal to 2.5%; *(iii)* the ratio of the 90th percentile to the median of the cross-sectional idiosyncratic volatility of stock returns, equal to 3.54; *(iv)* the average monthly equity turnover (defined as the total monthly volume divided by the number of shares outstanding), equal to 9.7%; *(v)* the fraction of assets that investors learn about, which, in the absence of

²⁶Specifically, we set $\sigma_i = \sigma_n + \alpha(n-i)/n$, which implies the volatility distribution is parameterized fully by a single parameter α .

²⁷Changing the number of assets in the parameterization does not have a major impact on our results, as long as the model is reparameterized to meet the same empirical targets.

²⁸To compute the number, we first compute the dollar value of the risky part of the financial holdings of households (stocks, bonds, non-money market funds, and other financials) for each decile of the wealth distribution. Then, we compute the share of these risky assets held by the top decile.

empirical guidance, we arbitrarily set to 50%; and (*vi*) the average annualized stock market excess return of the sophisticated investors of 13.1% and unsophisticated investors of 9.1%, whose values are discussed in the previous section.²⁹

5.2 Dynamics of Capital Income Inequality

We now assess our model’s quantitative predictions for the evolution of capital income inequality in response to aggregate growth in information technology. We set the initial capacity to be equal to the benchmark parameterization value, and we simulate the model for 25 years, reflecting the number of years in the SCF. Along the simulation path, we pick the capacity growth to match the overall excess return on the market for the entire period of 7%. This gives a 5.1% growth rate in annual information capacity, which fits within the range of independent data estimates on the increase in the number of stocks actively analyzed by the financial industry (4% growth annually) or the number of analysts per stock in the financial industry (8% growth annually). The results from this exercise are presented in Table 2. The model generates a 42% rise in capital income inequality, compared to 87% in the data, which means that our mechanism explains about 48% of the variation in the data.

To provide sensitivity of our findings to the assumed growth rate of the aggregate capacity, in Table 2, we also include growth in capital income inequality for 4% and 8% growth rates in the aggregate capacity. Within that range, the model accounts for 27% to 69% of the rise in capital income inequality in the data.

The Importance of Heterogeneity How important is it for our quantitative results to relax the commonly used assumption that households have access to a single risky asset? The fifth row of Table 2 presents results from an alternative specification

²⁹We perform a detailed grid search over parameters until all the simulated moments are within a 10% distance from target. That gives sophisticated ownership within 0.7%, sophisticated and unsophisticated returns within 7%, ratio of volatilities within 2% and all other targets matched exactly.

Table 2: Capital income inequality growth: benchmark model and robustness.

	Growth in inequality relative to 1989	% of data
Data	87%	
Benchmark	42%	48
Benchmark 4% growth	24%	27
Benchmark 8% growth	60%	69
One asset	20%	23
Asymmetric growth	54%	62

of the model with only one risky asset. The difference between this specification, labeled as *One asset*, and the benchmark model quantifies the role of asset heterogeneity in driving capital income inequality.³⁰ The one-asset economy generates growth in capital income inequality that is less than half of the growth generated by the benchmark model, and only 23% of the data. Hence, asset heterogeneity plays a crucial role in driving capital income inequality in the model. It generates higher payoffs from learning and larger effects on the retrenchment of unsophisticated investors from risky asset markets.

Asymmetric Growth We also investigate the importance of a heterogeneous growth in capacity for the evolution of inequality. Specifically, we assume that the growth in capacity of each investor type is proportional to that investor group’s re-

³⁰In terms of the parameterization, the model with one risky asset takes away three targets from the benchmark model: heterogeneity in asset volatility, fraction of actively traded assets, and the return of sophisticated investors. We keep the value of the risk aversion coefficient the same as in the benchmark model and pick the volatility of the single asset payoff equal to the median payoff volatility of the benchmark model (equal to 1.295). That leaves three remaining parameters: volatility of the noise trader demand σ_x , and the two capacities of sophisticated and unsophisticated investors. We choose these to match: the average market (also equal to sophisticated and unsophisticated) return (11.9%), average asset turnover (9.7%), and sophisticated ownership (69%). That gives $(K_1, K_2, \sigma_x) = (0.0544, 0.0163, 0.37)$. In simulating the model, we pick the growth rate of aggregate capacity just like in the benchmark model—to match the market return of 7% over the entire period. That implies the growth rate of technology of 6.7%, which is actually higher than the one implied by our benchmark specification.

turns, rather than to the market return on equity.³¹ This specification amplifies the feedback loop from high capacity to high returns, and hence increases the growth in inequality over time. As the last row of table Table 2 shows, the asymmetric growth model accounts for roughly 62% of the growth in inequality versus 48% in the benchmark model, which indicates an additional 30% effect due to asymmetric growth.

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, or skill. 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 $j \in \{S, U\}$:

$$R_j = E \sum_i \omega_{jit}(r_{it} - r) = \sum_i Cov(\omega_{jit}, r_{it}) + \sum_i E\omega_{jit}E[r_{it} - r], \quad (20)$$

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

³¹We also scale the constant of proportionality to be 0.93 in order for this exercise to exhibit the same average growth of aggregate capacity as does the benchmark model, equal to 5.1% annually.

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

$$\hat{R}_I = \sum_i Cov(\omega_{Rit}, r_{it}) + \sum_i E\omega_{Iit}E[r_{it} - r]. \quad (21)$$

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 2.4 percentage point differential between the sophisticated and unsophisticated investors.

Heterogeneity in Risk Aversion The overall growth in inequality can be increased by augmenting the model with differences in risk attitudes. In particular, if one group of investors were less risk averse they would hold a greater share of risky assets, and hence they would have higher expected capital income.³² Within our mean-variance specification, a growing difference in risk aversion produces growing *aggregate* ownership in risky assets of less risk averse investors, and a uniform, proportional retrenchment from risky assets of more risk averse investors. However, heterogeneity in risk aversion alone cannot generate the empirical investor-specific rates of return on equity, differences in portfolio weights within a class of risky assets or differential growth in ownership by asset volatility (discussed in the next section)./footnoteOn the other hand, Gomez (2016) shows that when macro- asset pricing models with heterogenous risk aversion are parameterized to match the volatility of asset prices, they require a degree of heterogeneity in preferences that leads to counterfactual predictions about wealth inequality. Hence, the information asymmetry would have to be retained.

³²Such setting would also encompass situations in which investors are exposed to different levels of volatility in areas outside capital markets, like labor income.

Alternative Preferences In the Online Appendix, we analyze the model with CRRA utility. Since a closed-form solution to the full model is not feasible, we focus on a local approximation of the utility function. We show that the model solution under no capacity differences predicts the same portfolio shares for risky assets, *independent of wealth*. Intuitively, if agents have common information, then wealth differences affect the composition of their allocations between the risk-free asset and the risky portfolio, but not the composition of the risky portfolio, which is determined optimally by the (common) belief structure. As a result, differences in capacity are a necessary component for the model to generate any risky return differences across agents.

Endogenous Capacity Choice In the benchmark model, we assume an exogenous relationship between initial capacity and an investor’s wealth. In the Online Appendix, we show how such relation could arise endogenously. Intuitively, if investors endogenously choose different portfolio sizes, then their net benefit of investing in information will increase with portfolio size. We apply this idea in a model in which investors have identical CRRA preferences and make endogenous capacity choice decisions. In the context of the information choice model, CRRA utility specification is not tractable; hence, we map a common relative risk aversion together with wealth differences locally into different absolute risk aversion coefficients. In a numerical example, we show how initial wealth differences observed in the 1989 SCF map into endogenous capacity differences, for different values of the cost of capacity and different relative risk aversion coefficients. We show that for a wide range of the risk aversion specifications and for capacity cost away from zero, the implied differences in capacity are equal or actually larger than the ones specified in the benchmark model. Hence, we view our parameterization as cautious in that it implies modest initial capacity differences.

5.3 Dynamic Predictions and External Validity

In this section, we generate a set of dynamic predictions of the model and compare them to the corresponding data moments to provide support for our mechanism. These are robust implications of our mechanism proven analytically in Section 3. Below, we show a quantitative fit of these predictions vis a vis the data.

We explore the consequences of the aggregate growth in capacity behind the results in Section 5.2. We compute statistics for the value of capacity that matches the market excess return for the second half of our sample, and relate them to 2001-2012 means in the data.

Market Averages In the model, symmetric growth in information capacities implies large changes in average market returns, cross-sectional return differentials, and turnover. Table 3 reports the model predictions and their empirical counterparts.

Table 3: Market Averages: Data and Model

Statistic	2001-2012	
	Data	Model
Market Returns	2.4%	2.4%
Sophisticated portfolio	2.9%	2.5%
Unsophisticated + Noise traders portfolio	1.1%	2.2%
Average Equity Turnover	16.0%	16.8%

Both the model and the data exhibit a decrease in market return and in the return difference between sophisticated and unsophisticated investors. If anything, the model actually underpredicts the difference in rates of return in the second half of the sample, which makes our prediction of accounting for 48% of the growth in capital income inequality a conservative one. The lower market return is a result of an increase in the quantity of information, as prices track payoffs more closely than in the initial sample period, implying lower excess returns. The model also predicts

a sharp increase in average asset turnover, in magnitudes consistent with the data. As with the market return, this result is a direct implication of our mechanism and is not driven by fundamental asset volatilities, which remain unchanged. Intuitively, higher turnover is driven by more informed trading by sophisticated investors, due to their holding a larger share of the market and receiving more precise signals about asset payoffs (Proposition 5).

Cross-sectional Turnover Our model implies a rich cross-sectional variation in asset turnover. Intuitively, if an asset is more attractive and investors want to trade it, then more investors with precise signals about this asset’s returns would want to act on such better information by taking larger and more volatile positions. Since sophisticated investors receive more precise signals, and they have preference for high-volatility assets, we should see a positive relationship between volatility and turnover. Table 4 reports turnover in relation to return volatility in the model and the data.

Table 4: Turnover by Asset Volatility

Volatility quintile	1	2	3	4	5	Mean
1989-2000						
Data	5%	8.5%	10.5%	12.5%	11.5%	9.7%
Model	8.9%	9.1%	9.4%	10.3%	11%	9.7%
2001-2012						
Data	11%	14.6%	17%	18.4%	19.3%	16%
Model	15.5%	16.7%	17.2%	17.3%	17.2%	16.8%

The first two rows compare data and the model predictions for the 1989-2000 sub-sample. Both data and model show that turnover is increasing in volatility, and the model’s predictions are quantitatively close to the data. In the next two rows, we compare data for the 2001-2012 period to results generated from the dynamic exercise. The model implies an increase in average turnover and additionally matches the cross-sectional pattern of this increase. This effect is purely driven by our information

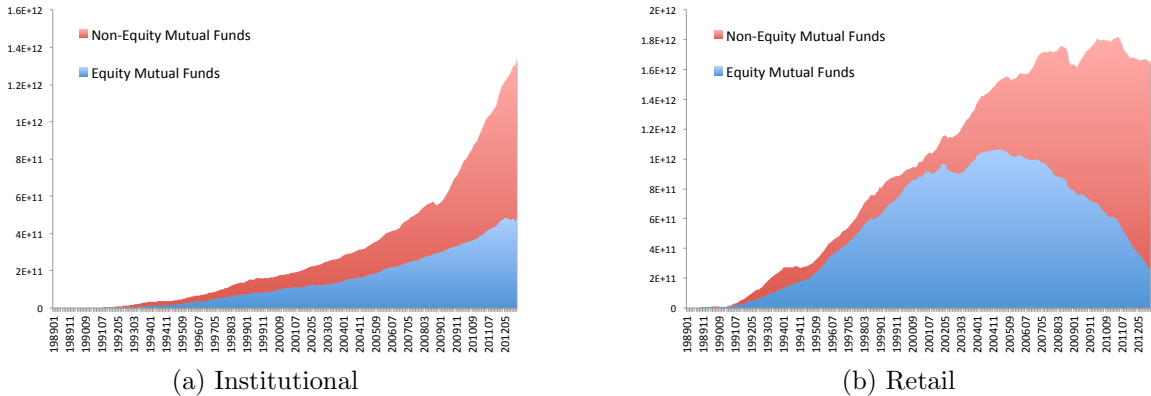


Figure 4: Cumulative Flows to Mutual Funds: Institutional vs. Retail

friction, since the fundamental volatilities in this exercise remain constant over time.³³

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. This result is summarized in Lemma 4.

To provide auxiliary empirical support in favor of the model’s ownership prediction, we consider flows into mutual funds. 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. The Morningstar data contains information for two types of funds: those with a minimum investment of \$100,000 (institutional funds) and those without such a threshold (retail funds). As shown in Figure 4, 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

³³Our model also implies a positive turnover-ownership relationship, which we confirm in the data. This result is consistent with the empirical findings in Chordia, Roll, and Subrahmanyam (2011).

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.

5.4 Mutual Funds and Delegation

In the data, households can invest in financial markets either directly or through intermediaries such as mutual funds. One could argue that unsophisticated households could bypass their information disadvantage by simply delegating their money to more informed asset management companies. But is it the case that any investor, regardless of wealth, can access the same quality of investment advice or intermediation? The previous section already discussed the barriers to accessing institutional mutual funds, which outperform retail funds by a significant margin. In this section, we take up the question of delegation in more detail.

First, we document a large heterogeneity in mutual fund returns in the data depending on the investment size (typically related to wealth). Additionally, 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.³⁴

³⁴We are not the first ones to point out these regularities. Extant literature in finance, such as Kacperczyk, Sialm, and Zheng (2005) or Pástor, Stambaugh, and Taylor (2015) finds that while the average abnormal gross returns of mutual funds are positive, the distribution of returns is highly

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 5 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.

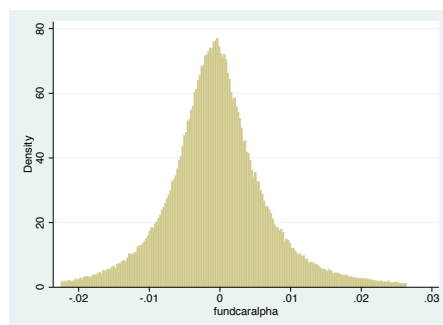


Figure 5: Distribution of equity funds' returns.

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. 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 dispersed and the returns are not predictable.

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 5.

Table 5: 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.

6 Concluding Remarks

What contributes to the growing income inequality across households? This question has been of great economic and policy relevance for at least several decades starting with the seminal work by Kuznets (1953). We approach this question from

the perspective of capital income that is known to be highly unequally distributed across individuals. We propose a theoretical information-based framework that links capital income derived from financial markets to a level of investor sophistication. Our model implies the presence of income inequality between sophisticated and unsophisticated investors that is growing in the extent of total sophistication in the market, and could be the result of aggregate technological progress. Additional predictions on asset ownership, market returns, and turnover help us pin down the economic mechanism and rule out alternative explanations. The quantitative predictions of the model match qualitatively and quantitatively the observed data.

One could argue that the overall growth of investment resources and competition across investors with different skill levels are generally considered as a positive aspect of a well-functioning financial market. However, our work suggests that one should assess any policy targeting overall information environment in financial markets as potentially exerting an offsetting and negative effect on socially relevant issues, such as distribution of income. Our work also sheds light on the overall benefits and redistribution aspects of progress in financial markets in terms of creating new financial instruments. Depending on where the new assets land on the volatility (or more generally, opaqueness) spectrum, the benefits will accrue to the relatively less (low-volatility assets) or more (high-volatility assets) sophisticated investors.

References

- Aiyagari, Rao S., 1994, Uninsured idiosyncratic risk and aggregate saving, *Quarterly Journal of Economics* 109(3), 659–684.
- Arrow, Kenneth J, 1987, The demand for information and the distribution of income, *Probability in the Engineering and Informational Sciences* 1, 3–13.
- Atkinson, Anthony B, Thomas Piketty, and Emmanuel Saez, 2011, Top incomes over a century or more, *Journal of Economic Literature* 49, 3–71.

- Bach, Laurent, Laurent Calvet, and Paolo Sodini, 2015, Rich pickings? Risk, return, and skill in the portfolios of the wealthy, Working Paper Stockholm School of Economics.
- Benhabib, Jess, and Alberto Bisin, 2017, Skewed wealth distributions: Theory and empirics, *Journal of Economic Literature* forthcoming.
- Benhabib, Jess, Alberto Bisin, and Shenghao Zhu, 2011, The distribution of wealth and fiscal policy in economies with finitely lived agents, *Econometrica* 79, 123–157.
- Cagetti, Marco, and Mariacristina De Nardi, 2006, Entrepreneurship, frictions, and wealth, *Journal of Political Economy* 114, 835–870.
- Calvet, Laurent E, John Y Campbell, and Paolo Sodini, 2009, Measuring the financial sophistication of households, *American Economic Review* 99, 393–398.
- Chien, YiLi, Harold Cole, and Hanno Lustig, 2011, A multiplier approach to understanding the macro implications of household finance, *Review of Economic Studies* 78 (1), 199–234.
- Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2011, Recent trends in trading activity and market quality, *Journal of Financial Economics* 101, 243–263.
- De Nardi, Mariacristina, and Giulio Fella, 2017, Saving and wealth inequality, CEPR Working Paper 11746.
- Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, 2016a, Heterogeneity and persistence in returns to wealth, NBER Working Paper No 22822.
- Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, 2016b, Heterogeneity in returns to wealth and the measurement of wealth inequality, *The American Economic Review* 106, 651–655.
- Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, 2016, The dynamics of inequality, *Econometrica* 84, 2071–2111.
- Gomez, Matthieu, 2016, Asset prices and wealth inequality, Working Paper Princeton University.
- Greenwood, Daphne, 1983, An estimation of US family wealth and its distribution from microdata, *Review of Income and Wealth* 29 (1), 23–44.
- Guvenen, Fatih, 2007, Do stockholders share risk more effectively than non-stockholders?, *Review of Economics and Statistics* 89(2), 275–288.
- Guvenen, Fatih, 2009, A parsimonious macroeconomic model for asset pricing, *Econometrica* 77(6), 1711–1750.

- Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2012.
- Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2016, A rational theory of mutual funds’ attention allocation, *Econometrica* 84(2), 571–626.
- Kessler, Dennis, and Edward N. Wolff, 1991, A comparative analysis of household wealth patterns in France and the United States, *Review of Income and Wealth* 37 (1), 249–266.
- Krusell, Per, and Anthony A. Smith, 1998, Income and wealth heterogeneity in the macroeconomy, *Journal of Political Economy* 106(5), 867–896.
- Kuznets, Simon, 1953, Shares of Upper Income Groups in Savings, in *Shares of Upper Income Groups in Income and Savings* (NBER, New York).
- Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S. Mitchell, 2017, Optimal financial knowledge and wealth inequality, *Journal of Political Economy* forthcoming.
- Mankiw, N Gregory, and Stephen P Zeldes, 1991, The consumption of stockholders and nonstockholders, *Journal of Financial Economics* 29, 97–112.
- Pástor, Luboš, Robert F. Stambaugh, and Luke Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.
- Pástor, Luboš, and Pietro Veronesi, 2016, Income inequality and asset prices under redistributive taxation, *Journal of Monetary Economics* 81, 1–20.
- Peress, Joel, 2004, Wealth, information acquisition and portfolio choice, *The Review of Financial Studies* 17(3), 879–914.
- Piketty, Thomas, 2014, *Capital in the Twenty-First Century*. (Harvard University Press).
- Piketty, Thomas, and Emmanuel Saez, 2003, Income inequality in the United States, 1913–1998, *The Quarterly Journal of Economics* 118, 1–41.
- Quadrini, Vincenzo, 1999, The importance of entrepreneurship for wealth concentration and mobility, *Review of Income and Wealth* 45(1), 1–19.
- Saez, Emmanuel, and Gabriel Zucman, 2016, Wealth inequality in the United States since 1913: Evidence from capitalized income tax data, *The Quarterly Journal of Economics* 131, 519–578.
- Shannon, Claude E, 1948, A mathematical theory of communication, *Bell System Technical Journal* 27, 379–423 and 623–656.

- Sims, Christopher A, 2003, Implications of rational inattention, *Journal of Monetary Economics* 50(3), 665–690.
- Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2009, Information immobility and the home bias puzzle, *Journal of Finance* 64(3), 1187–1215.
- Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2010, Information acquisition and under-diversification, *Review of Economic Studies* 77(2), 779–805.
- Vissing-Jorgensen, Annette, 2004, Perspectives on behavioral finance: Does “irrationality” disappear with wealth? Evidence from expectations and actions, in *NBER Macroeconomics Annual 2003, Vol. 18* (The MIT Press, Cambridge).

Appendix: Proofs

Model

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

For each asset i , the entropy of $z_i \sim \mathcal{N}(\bar{z}, \sigma_i^2)$ 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

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

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

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

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

Information Objective. Expected utility is given by

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

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

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

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

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

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 $\text{Cov}(\hat{\mu}_{ji}, p_i)$ exploiting the fact that posterior beliefs and prices are conditionally independent given payoffs. We obtain

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

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

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

Expected utility becomes

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

where $G_i \equiv (1 - rb_i)^2 + \frac{r^2 c_i^2 \sigma_x^2}{\sigma_i^2} + \frac{(\bar{z} - ra_i)^2}{\sigma_i^2}$, 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. For all other assets, the posterior variance is equal to the prior variance. Let l_j index the asset about which investor j learns. The information constraint becomes $\prod_{i=1}^n \frac{\sigma_i^2}{\hat{\sigma}_{ji}^2} = \frac{\sigma_{l_j}^2}{\hat{\sigma}_{j l_j}^2} = e^{2K_j}$, and hence the variance of the investor's beliefs is given by

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

The investor's problem becomes picking the asset l_j to maximize $\sum_{i=1}^n G_i \frac{\sigma_i^2}{\hat{\sigma}_{ji}^2} = (e^{2K_j} - 1) G_{l_j} + \sum_{i=1}^n G_i$. 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 distribution of posterior beliefs follows. \square

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

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

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

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

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

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

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

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

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

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

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

Proof of Lemma 3 (Equilibrium Learning). Substituting a_i , b_i , and c_i in $G_i \equiv (1 - rb_i)^2 + \frac{r^2 c_i^2 \sigma_x^2}{\sigma_i^2} + \frac{(\bar{z} - ra_i)^2}{\sigma_i^2}$ and defining $\xi_i \equiv \sigma_i^2 (\sigma_x^2 + \bar{x}^2)$ 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 $i \in \{1, \dots, n-1\}$. 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 $i \in \{2, \dots, k\}$, 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 $i \in \{k+1, \dots, n\}$. Recursively, $m_i = c_{i1} m_1 - \frac{1}{\phi} (1 - c_{i1})$, $\forall i \in \{2, \dots, k\}$. 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 , $i \in \{1, \dots, k\}$, $m_i = \frac{c_{i1}}{C_k} + \frac{1}{\phi} \left(\frac{k c_{i1}}{C_k} - 1 \right)$. \square

Proof of Lemma 4 (Learning Dynamics). (i) First, consider a local increase in ϕ to some $\phi' \leq \phi_k$, such that no new assets are learned about in equilibrium (k and C_k are unchanged). For $i \in \{1, \dots, k\}$,

$$\frac{dm_i}{d\phi} = -\frac{1}{\phi^2} \left(\frac{k c_{i1}}{C_k} - 1 \right), \text{ where } c_{i1} \equiv \sqrt{\frac{1 + \rho^2 \xi_i}{1 + \rho^2 \xi_1}} \leq 1 \text{ and } C_k \equiv \sum_{i=1}^k c_{i1}.$$

Hence m_i is strictly decreasing in ϕ if $c_{i1} > \frac{C_k}{k}$ (namely, if the asset is above average in terms of relative volatility), and m_i is increasing in ϕ otherwise. Since c_{i1} is decreasing in i , the condition $c_{\bar{i}1} = C_k/k$ defines the cutoff asset \bar{i} . Moreover, note that for $i \in \{1, \dots, \bar{i}\}$, the absolute value of $\frac{dm_i}{d\phi}$ is decreasing in i , such that the masses of the more volatile assets fall by more than those of the less volatile assets. Likewise, for $i \in \{\bar{i} + 1, \dots, k\}$, the value of $\frac{dm_i}{d\phi}$ is increasing in i , such that the masses of the less volatile assets increase by more than those of the more volatile assets. This results in a flattening of the distribution of investors across assets.

Next, suppose that $k < n$, and consider an increase in ϕ to some $\phi' > \phi_k$, such that $k' > k$ assets are learned about (with $k' \leq n$). Let the new equilibrium masses be denoted by m'_i

for $i \in \{1, \dots, k'\}$. Hence, $\sum_{i=1}^k m'_i < 1$. Using the recursive expression for m_i in terms of m_1 , for $i \in \{2, \dots, k\}$

$$m_i - m'_i = c_{i1} (m_1 - m'_1) - (1 - c_{i1}) \left(\frac{1}{\phi} - \frac{1}{\phi'} \right).$$

Suppose that $m_1 \leq m'_1$. Then $\sum_{i=1}^k m_i - \sum_{i=1}^k m'_i = 1 - \sum_{i=1}^k m'_i < 0$, which is a contradiction. Hence $m_1 > m'_1$. Moreover, since c_{i1} is decreasing in i , the condition $m_{\bar{i}} = m'_{\bar{i}}$ defines the threshold value for c_{i1} that defines the cutoff asset \bar{i} .

(ii) First, consider a local increase in ϕ to some $\phi' < \phi_k$, such that no new assets are learned about (k and C_k are unchanged). For $i \in \{1, \dots, k\}$,

$$\frac{d(\phi m_i)}{d\phi} = \frac{c_{i1}}{C_k} > 0.$$

Next, suppose that $k < n$, and consider an increase in ϕ to some $\phi' > \phi_k$, such that $k' > k$ assets are learned about in equilibrium (with $k' \leq n$). First, for the new assets that are actively traded, $i \in \{k+1, \dots, k'\}$, $m'_i > m_i = 0$, hence, $\phi' m'_i > \phi m_i$. Second, consider an asset $i \in \{1, \dots, k\}$ and an asset $h \in \{k+1, \dots, k'\}$. Let the new equilibrium gains be denoted by G'_i and G'_h . Then $G_i > G_h$, which implies that $1 + \phi m_i < c_{ih}$, and $G'_i = G'_h$, which implies that $1 + \phi' m'_i = (1 + \phi' m'_h) c_{ih} > (1 + \phi' m'_h) (1 + \phi m_i) \Leftrightarrow \phi' m'_i > \phi m_i + \phi' m'_h (1 + \phi m_i) > \phi m_i$.

(iii) Let $K_1 = K$ and $K_2 = \delta K$, for some $\delta \in (0, 1)$, and consider a symmetric increase in K_j to $(1 + \gamma)K_j$ such that the first $k' \geq k$ assets are learned about.

Let \bar{i} denote the cutoff asset determined in part (i). For $i \in \{\bar{i}, \dots, k'\}$, both $m_i (e^{2K_1} - 1)$ and $m_i (e^{2K_2} - 1)$ increase, since $m'_i \geq m_i$, but $m_i (e^{2K_1} - 1)$ grows by more since e^x is convex.

For $i \in \{1, \dots, \bar{i} - 1\}$, m_i is decreasing in ϕ .

Let $m_{i\phi} \equiv \frac{dm_i}{d\phi}$. The derivatives of interest are

$$D_1 \equiv \frac{d[m_i(e^{2K_1}-1)]}{dK} = m_{i\phi} (e^{2K} - 1) \frac{d\phi}{dK} + 2e^{2K} m_i$$

$$D_2 \equiv \frac{d[m_i(e^{2K_2}-1)]}{dK} = m_{i\phi} (e^{2K\delta} - 1) \frac{d\phi}{dK} + 2\delta e^{2K\delta} m_i$$

where $\frac{d\phi}{dK} = 2\lambda e^{2K} + 2\delta(1 - \lambda) e^{2K\delta} > 0$.

Factoring out $2e^{2K}$ yields

$$\begin{aligned} D_1 &= 2e^{2K} \{m_i + m_{i\phi} (e^{2K} - 1) [\lambda + (1 - \lambda) \delta e^{2K(\delta-1)}]\} \\ &= 2e^{2K} \{m_i + m_{i\phi} [\lambda (e^{2K} - 1) + (1 - \lambda) \delta (e^{2K\delta} - e^{2K(\delta-1)})]\} \\ &\stackrel{(1)}{>} 2e^{2K} \{m_i + m_{i\phi} [\lambda (e^{2K} - 1) + (1 - \lambda) (e^{2K\delta} - 1)]\} \\ &= 2e^{2K} \{m_i + m_{i\phi} \phi\} = 2e^{2K} \left[\frac{d(\phi m_i)}{d\phi} \right] \stackrel{(2)}{>} 0, \end{aligned}$$

where (2) follows from part (ii) above and (1) follows from the evaluation of the function

$F(\delta) = e^{2K\delta} - 1 - \delta(e^{2K\delta} - e^{2K(\delta-1)})$ for which $F(0) = F(1) = 0$, $F'(\delta) = 0$ has a unique solution, $F'(\delta) > 0$ and $F'(\delta) < 0$, which imply that $F(\delta) > 0$.

Next, note that $\lambda D_1 + (1 - \lambda) D_2 = \left[\frac{d(\phi m_i)}{d\phi} \right] \frac{d\phi}{dK} = 2 \left[\frac{d(\phi m_i)}{d\phi} \right] [\lambda e^{2K} + \delta(1 - \lambda) e^{2K\delta}]$. We have just shown that $D_1 > 2e^{2K} \left[\frac{d(\phi m_i)}{d\phi} \right]$, so for the equality to hold it must be the case that $D_2 < 2\delta e^{2K\delta} \left[\frac{d(\phi m_i)}{d\phi} \right]$. Hence $D_1 > 0$ and $D_1 > D_2$. It remains to be determined if $D_2 > 0$ as well.

$$\begin{aligned} D_2 &= 2e^{2K\delta} \left\{ \delta m_i + m_{i\phi} (1 - e^{-2K\delta}) [\lambda e^{2K} + \delta(1 - \lambda) e^{2K\delta}] \right\} \\ &= 2e^{2K\delta} \left\{ \delta m_i + m_{i\phi} [\lambda e^{2K} + \delta(1 - \lambda) e^{2K\delta} - \lambda e^{2K-2K\delta} - \delta(1 - \lambda)] \right\} \\ &= 2e^{2K\delta} \left\{ \delta m_i + m_{i\phi} [\lambda (e^{2K} - 1) - \lambda (e^{2K-2K\delta} - 1) + \delta(1 - \lambda) (e^{2K\delta} - 1)] \right\} \\ &> 2e^{2K\delta} \left\{ \delta m_i + m_{i\phi} [\lambda (e^{2K} - 1) + \delta(1 - \lambda) (e^{2K\delta} - 1)] \right\} \\ &> 2e^{2K\delta} \left\{ \delta m_i + m_{i\phi} [\lambda (e^{2K} - 1) + (1 - \lambda) (e^{2K\delta} - 1)] \right\} \\ &= 2e^{2K\delta} \left\{ \delta m_i + m_{i\phi} \phi \right\} = 2e^{2K\delta} \left[\frac{d(\phi m_i)}{d\phi} \right] - (1 - \delta) m_i. \end{aligned}$$

Hence, if δ is not too small (i.e. capacity dispersion is not too large), then $D_2 > 0$ for $i \in \{1, \dots, \bar{i} - 1\}$ as well.

Hence, for assets $i \in \{1, \dots, \bar{i} - 1\}$, for which the mass of investors falls in response to the capacity growth, $m_i (e^{2K_1} - 1)$ grows and $m_i (e^{2K_2} - 1)$ grows by less, or even falls, if capacity dispersion is large enough. \square

Analytic Results

Proof of Proposition 1. Results follow from equations (13-15). \square

Proof of Proposition 2. (i) Follows from the definition of capital income per capita and equation (14). (ii) Since for all $i \in \{1, \dots, k\}$, 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, using equation (10), the masses m_i are unchanged. With both ϕ and m_i unchanged, prices are unchanged. (ii) The result follows from equation (14): masses and prices do not change, and dispersion, $(e^{2K_1} - e^{2K_2})$ increases. (iii) Relative capital income is

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

Since prices are unchanged, $(\bar{z}_i - rp_i)(z_i - rp_i)$ 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, $\bar{p}_i = \frac{1}{r} \left(\bar{z} - \frac{\rho\sigma_i^2\bar{x}}{1+\phi m_i} \right)$. Per Lemma 4, ϕm_i is increasing in ϕ . Hence, for $i \in \{1, \dots, k\}$, \bar{p}_i is increasing in ϕ . The result for equilibrium expected excess returns $E[z_i - r\bar{p}_i]$ follows.

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

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

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

In the proof of Lemma 4, we show that $\frac{d[m_i(e^{2K}-1)]}{dK} > 2e^{2K} \frac{d(\phi m_i)}{d\phi} > 0$. Hence,

$$\begin{aligned} \text{sign} \left(\frac{dE[q_{1i}]}{dK} \right) &= \text{sign} \left(2e^{2K} - \frac{d\phi}{dK} \frac{m_i(e^{2K}-1)}{1+\phi m_i} \right) \\ &= \text{sign} \left(2e^{2K} - \frac{2m_i[\lambda e^{2K} + (1-\lambda)\delta e^{2K\delta}](e^{2K}-1)}{1+m_i[\lambda(e^{2K}-1) + (1-\lambda)(e^{2K\delta}-1)]} \right) \\ &= \text{sign} \left(e^{2K} - (e^{2K}-1) \frac{m_i[\lambda e^{2K} + (1-\lambda)\delta e^{2K\delta}]}{1+m_i[\lambda e^{2K} + (1-\lambda)e^{2K\delta}] - m_i} \right) \\ &\stackrel{(1)}{=} \text{sign} \left(e^{2K} - (e^{2K}-1) \left[\frac{m_i[\lambda e^{2K} + (1-\lambda)e^{2K\delta}]}{1+m_i[\lambda e^{2K} + (1-\lambda)e^{2K\delta}] - m_i} \right] \right) \\ &\stackrel{(2)}{=} \text{sign} (e^{2K} - (e^{2K}-1)) > 0 \end{aligned}$$

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:

$\pi_{1i} = c_i + \frac{1}{\rho\sigma_i^2} m_i (e^{2K} - 1) (z_i - rp_i)^2$, where $c_i \equiv \frac{1}{\rho\sigma_i^2} (\bar{z} - rp_i) (z_i - rp_i)$, with expected value C_i . Then $E[\pi_{1i}] = C_i + \frac{1}{\rho\sigma_i^2} m_i (e^{2K} - 1) E[(z_i - rp_i)^2] = C_i + \frac{1}{\rho} m_i (e^{2K} - 1) G_i$, where G_i is the gain from learning about asset i , equated across all $i \in \{1, \dots, k\}$.

We then obtain that $\frac{E[\pi_{1i}]}{E[\pi_{2i}]} = \frac{C_i + \frac{1}{\rho} m_i (e^{2K}-1) G_i}{C_i + \frac{1}{\rho} m_i (e^{2K\delta}-1) G_i}$.

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. From Lemma 4, $m_i (e^{2K} - 1)$ increases by more than $m_i (e^{2K\delta} - 1)$ in response to a change in K . Hence overall, $\frac{E[\pi_{1i}]}{E[\pi_{2i}]}$ increases. \square

Derivation of volume per capita. We define the volume of trade in asset i between two periods, across all optimizing investors j in group g as $V_i^g \equiv \int |q'_{ji} - q_{ji}| dj$. Integrating over all possible realizations of q'_{ji} and q_{ji} , we obtain average volume across many periods, \bar{V}_i^g . We assume that investors do not change groups over time. To ease notation, most of the derivation omits group and asset superscripts.

Volume between two periods for a generic group First, we calculate the expected volume of trade for each asset by agents in each group from period t to $t + 1$. Let f and F denote the pdf and cdf of current holdings, with mean μ_q and standard deviation σ_q . Let f' and F' denote the pdf and cdf of future holdings, with mean μ'_q and standard deviation σ'_q .

STEP 1. Consider a particular investor with holdings q in the current period. The investor's expected volume of trade between the current and the next period is

$$v(q) \equiv \int_{-\infty}^{+\infty} |q' - q| f'(q') dq' = 2qF(q) - q - 2F'(q) E_{f'}[q' | q' < q] + \mu'_q.$$

Using the formula for the expected value of a normal truncated from above,

$$v(q) = 2qF(q) - q - 2\mu'_q F'(q) + 2\sigma_q'^2 f'(q) + \mu'_q.$$

STEP 2. Integrating over the (normal) distribution of holdings q in the group,

$$V^g = \int_{-\infty}^{+\infty} v(q) f(q) dq = 2 \int_{-\infty}^{+\infty} qF(q) f(q) dq - \mu_q - 2\mu'_q \int_{-\infty}^{+\infty} F'(q) f(q) dq + 2\sigma_q'^2 \int_{-\infty}^{+\infty} f'(q) f(q) dq + \mu'_q.$$

Using the formulas $\int_{-\infty}^{+\infty} \exp\{-ax^2 + bx + c\} dx = \sqrt{\frac{\pi}{a}} \exp\left\{\frac{b^2}{4a} + c\right\}$,

$$\int_{-\infty}^{+\infty} \Phi(a + bx) \phi(x) dx = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right) \text{ and } \int_{-\infty}^{+\infty} x\Phi(bx) \phi(x) dx = \frac{b}{\sqrt{2\pi(1+b^2)}},$$

we compute $J_1 \equiv \int_{-\infty}^{+\infty} qF(q) f(q) dq = \frac{\mu_q}{2} + \frac{\sigma_q}{2\sqrt{\pi}}$,

$$J_2 \equiv \int_{-\infty}^{+\infty} F'(q) f(q) dq = \Phi\left(\frac{\mu_q - \mu'_q}{\sqrt{\sigma_q^2 + \sigma_q'^2}}\right),$$

$$J_3 \equiv \int_{-\infty}^{+\infty} f'(q) f(q) dq = \frac{1}{\sqrt{2\pi(\sigma_q^2 + \sigma_q'^2)}} \exp\left\{-\frac{(\mu_q - \mu'_q)^2}{2(\sigma_q^2 + \sigma_q'^2)}\right\}.$$

Hence $V^g = \frac{\sigma_q}{\sqrt{\pi}} - 2\mu'_q \Phi\left(\frac{\mu_q - \mu'_q}{\sqrt{\sigma_q^2 + \sigma_q'^2}}\right) + \frac{2\sigma_q'^2}{\sqrt{2\pi(\sigma_q^2 + \sigma_q'^2)}} \exp\left\{-\frac{(\mu_q - \mu'_q)^2}{2(\sigma_q^2 + \sigma_q'^2)}\right\} + \mu'_q$, where the means and standard deviations are group and asset specific.

Since the shocks are i.i.d., holdings have the same cross-sectional variance in all periods, $\sigma'_q = \sigma_q$, though they will have different means, depending on shock realizations. Hence

$$V^g = \frac{\sigma_q}{\sqrt{\pi}} \left[1 + \exp \left\{ -\frac{(\mu_q - \mu'_q)^2}{4\sigma_q^2} \right\} \right] + \mu'_q \left[1 - 2\Phi \left(\frac{\mu_q - \mu'_q}{\sqrt{2\sigma_q^2}} \right) \right].$$

Average volume across many periods for a generic group We assume no change in the environment, including no change in capacities and hence learning. Let the distribution of mean holdings μ_q be denoted by g , with mean and variance μ_μ and σ_μ^2 . Since the shocks are normal i.i.d., this distribution is stationary and also normal. The average volume across all possible realizations of μ_q and μ'_q is

$$\bar{V}^g = \int_{-\infty}^{+\infty} v^g(\mu_q) g(\mu_q) d\mu_q, \text{ with } v^g(\mu_q) = \int_{-\infty}^{+\infty} V^g(\mu_q, \mu'_q) g(\mu'_q) d\mu'_q.$$

STEP 3. Using the expression for V^g ,

$$v^g(\mu_q) = \frac{\sigma_q}{\sqrt{\pi}} + \frac{\sigma_q}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \exp \left\{ -\frac{(\mu_q - \mu'_q)^2}{4\sigma_q^2} \right\} g(\mu'_q) d\mu'_q + \mu_\mu - 2 \int_{-\infty}^{+\infty} \mu'_q \Phi \left(\frac{\mu_q - \mu'_q}{\sqrt{2\sigma_q^2}} \right) g(\mu'_q) d\mu'_q$$

Using the formulas for integrals of normal distributions, we compute

$$J_1 \equiv \int_{-\infty}^{+\infty} \exp \left\{ -\frac{(\mu_q - \mu'_q)^2}{4\sigma_q^2} \right\} g(\mu'_q) d\mu'_q = \sqrt{\frac{2\sigma_q^2}{\sigma_\mu^2 + 2\sigma_q^2}} \exp \left\{ -\frac{(\mu_q - \mu_\mu)^2}{2(\sigma_\mu^2 + 2\sigma_q^2)} \right\},$$

$$J_2 \equiv \int_{-\infty}^{+\infty} \mu'_q \Phi \left(\frac{\mu_q - \mu'_q}{\sqrt{2\sigma_q^2}} \right) g(\mu'_q) d\mu'_q = \mu_\mu \Phi \left(\frac{\mu_q - \mu_\mu}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \right) - \frac{\sigma_\mu^2}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \phi \left(\frac{\mu_q - \mu_\mu}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \right).$$

Then $v^g(\mu_q) = \frac{\sigma_q}{\sqrt{\pi}} + \frac{\sigma_q}{\sqrt{\pi}} J_1 + \mu_\mu - 2J_2$ becomes

$$\frac{\sigma_q}{\sqrt{\pi}} + \frac{\sigma_q^2 \sqrt{2}}{\sqrt{\pi}(\sigma_\mu^2 + 2\sigma_q^2)} \exp \left\{ -\frac{(\mu_q - \mu_\mu)^2}{2(\sigma_\mu^2 + 2\sigma_q^2)} \right\} + \mu_\mu - 2\mu_\mu \Phi \left(\frac{\mu_q - \mu_\mu}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \right) + \frac{2\sigma_\mu^2}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \phi \left(\frac{\mu_q - \mu_\mu}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \right).$$

STEP 4. Finally, integrating $v^g(\mu_q)$ over all possible realizations of μ_q , we obtain

$$\begin{aligned} \bar{V}^g &= \frac{\sigma_q}{\sqrt{\pi}} + \frac{\sigma_q^2 \sqrt{2}}{\sqrt{\pi}(\sigma_\mu^2 + 2\sigma_q^2)} \int_{-\infty}^{+\infty} \exp \left\{ -\frac{(\mu_q - \mu_\mu)^2}{2(\sigma_\mu^2 + 2\sigma_q^2)} \right\} g(\mu_q) d\mu_q + \mu_\mu \\ &\quad - 2\mu_\mu \int_{-\infty}^{+\infty} \Phi \left(\frac{\mu_q - \mu_\mu}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \right) g(\mu_q) d\mu_q + \frac{2\sigma_\mu^2}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \int_{-\infty}^{+\infty} \phi \left(\frac{\mu_q - \mu_\mu}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \right) g(\mu_q) d\mu_q. \end{aligned}$$

We compute

$$J_1 \equiv \int_{-\infty}^{+\infty} \exp \left\{ -\frac{(\mu_q - \mu_\mu)^2}{2(\sigma_\mu^2 + 2\sigma_q^2)} \right\} g(\mu_q) d\mu_q = \sqrt{\frac{\sigma_\mu^2 + 2\sigma_q^2}{2(\sigma_\mu^2 + 2\sigma_q^2)}},$$

$$J_2 \equiv \int_{-\infty}^{+\infty} \Phi \left(\frac{\mu_q - \mu_\mu}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \right) g(\mu_q) d\mu_q = \frac{1}{2},$$

$$J_3 \equiv \int_{-\infty}^{+\infty} \phi \left(\frac{\mu_q - \mu_\mu}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} \right) g(\mu_q) d\mu_q = \frac{J_1}{\sqrt{2\pi}}.$$

Then

$$\bar{V}^g = \frac{\sigma_q}{\sqrt{\pi}} + \frac{\sigma_q^2 \sqrt{2}}{\sqrt{\pi(\sigma_\mu^2 + 2\sigma_q^2)}} J_1 + \mu_\mu - 2\mu_\mu J_2 + \frac{2\sigma_\mu^2}{\sqrt{\sigma_\mu^2 + 2\sigma_q^2}} J_3 = \frac{1}{\sqrt{\pi}} \left(\sigma_q + \sqrt{\sigma_q^2 + \sigma_\mu^2} \right).$$

Variations by Investor Group

Consider the groups $g = SL, UL$ of sophisticated and unsophisticated investors who learn about asset i . These groups differ in their capacities only. A particular investor j in group g holds $q_{ji} = e^{2K_g}(s_{ji} - rp_i)/\rho\sigma_i^2$. The cross-sectional variance of holdings for this group, conditional on the realized shocks, is

$$\left(\sigma_{qi}^g \right)^2 = \left(\frac{e^{4K_g}}{\rho^2 \sigma_i^4} \right) \text{Var}(s_{ji} - rp_i) = \frac{e^{2K_g} - 1}{\rho^2 \sigma_i^2}.$$

The cross-sectional mean is

$$\mu_{qi}^g = \left(\frac{e^{2K_g}}{\rho \sigma_i^2} \right) E(s_{ji} - rp_i) = \frac{e^{2K_g}}{1 + \phi m_i} (\bar{x} + \nu_i) + \frac{e^{2K_g} - 1 - \phi m_i}{\rho \sigma_i^2 (1 + \phi m_i)} \varepsilon_i.$$

The expected value of mean of holdings is $\mu_{\mu i}^g = \frac{e^{2K_g}}{1 + \phi m_i} \bar{x}$ and the variance of mean holdings is

$$\left(\sigma_{\mu i}^g \right)^2 = \left(\frac{e^{2K_g}}{1 + \phi m_i} \right)^2 \sigma_x^2 + \left(\frac{e^{2K_g} - 1 - \phi m_i}{1 + \phi m_i} \right)^2 \frac{1}{\rho^2 \sigma_i^2}.$$

Consider the group NL of investors who are not learning about asset i . All investors in this group hold the same quantity $q_{ji} = \mu_{qi} = (\bar{z} - rp_i)/\rho\sigma_i^2$. Hence

$$\left(\sigma_{qi}^{NL} \right)^2 = 0 \text{ and } \mu_{qi}^{NL} = \frac{1}{\rho \sigma_i^2} (\bar{z} - rp_i).$$

The mean and variance of mean holdings are

$$\mu_{\mu i}^{NL} = \frac{\bar{x}}{1 + \phi m_i} \text{ and } \left(\sigma_{\mu i}^{NL} \right)^2 = \left(\frac{1}{1 + \phi m_i} \right)^2 \sigma_x^2 + \left(\frac{\phi m_i}{1 + \phi m_i} \right)^2 \frac{1}{\rho^2 \sigma_i^2}.$$

Consider the assets with zero learning, ZL . For assets that are not learned about by anyone ($m_i = 0, \phi m_i = 0$), all investors hold $q_{ji} = \mu_{qi} = (\bar{z} - rp_i)/\rho\sigma_i^2$. Hence

$$\left(\sigma_{qi}^{ZL} \right)^2 = 0 \text{ and } \mu_{qi}^{ZL} = \frac{1}{\rho \sigma_i^2} (\bar{z} - rp_i).$$

The mean and variance of mean holdings are

$$\mu_{\mu i}^{ZL} = \bar{x} \text{ and } \left(\sigma_{\mu i}^{ZL} \right)^2 = \sigma_x^2.$$

□

Proof of Proposition 5. First, average volume of active investors, $g = SL, UL$, is

$$\bar{V}^g = \frac{1}{\sqrt{\pi\rho^2\sigma_i^2}} \left[\sqrt{e^{2K_g} - 1} + \sqrt{e^{2K_g} - 1 + \left(\frac{e^{2K_g}}{1+\phi m_i}\right)^2 \rho^2 \sigma_i^2 \sigma_x^2 + \left(\frac{e^{2K_g} - 1 - \phi m_i}{1+\phi m_i}\right)^2} \right]$$

\bar{V}^g is increasing in K_g hence $\bar{V}^{SL} > \bar{V}^{UL}$.

Next, average volume of passive investors in actively traded assets is

$$\bar{V}_i^{NL} = \frac{\sigma_\mu^{NL}}{\sqrt{\pi}} = \frac{1}{\sqrt{\pi\rho^2\sigma_i^2}} \sqrt{\left(\frac{\rho\sigma_i\sigma_x}{1+\phi m_i}\right)^2 + \left(\frac{\phi m_i}{1+\phi m_i}\right)^2}.$$

Using $\sqrt{a} + \sqrt{b} > \sqrt{a+b}$,

$$\begin{aligned} \bar{V}^{UL} &> \frac{1}{\sqrt{\pi\rho^2\sigma_i^2}} \left[\sqrt{2(e^{2K_2} - 1) + \left(\frac{e^{2K_2}}{1+\phi m_i}\right)^2 \rho^2 \sigma_i^2 \sigma_x^2 + \left(\frac{e^{2K_2} - 1 - \phi m_i}{1+\phi m_i}\right)^2} \right] \\ &= \frac{1}{\sqrt{\pi\rho^2\sigma_i^2}} \sqrt{\frac{(\phi m_i)^2(2e^{2K_2} - 1) + 2\phi m_i(e^{2K_2} - 1) + (e^{4K_2} - 1)}{(1+\phi m_i)^2} + \left(\frac{e^{2K_2}}{1+\phi m_i}\right)^2 \rho^2 \sigma_i^2 \sigma_x^2} \\ &> \frac{1}{\sqrt{\pi\rho^2\sigma_i^2}} \sqrt{\frac{(\phi m_i)^2}{(1+\phi m_i)^2} + \left(\frac{1}{1+\phi m_i}\right)^2 \rho^2 \sigma_i^2 \sigma_x^2} = \bar{V}_i^{NL}. \end{aligned}$$

Hence for $i \in \{1, \dots, k\}$, $\bar{V}_i^{SL} > \bar{V}_i^{UL} > \bar{V}_i^{NL}$.

□