

Internet Appendix for  
“How Competitive is the Stock Market?  
Theory, Evidence from Portfolios, and Implications for the Rise of  
Passive Investing”\*

Valentin Haddad  
UCLA and NBER

Paul Huebner  
UCLA

Erik Loualiche  
University of Minnesota

August 8, 2022

---

\*Haddad: [vhaddad@ad.ucla.edu](mailto:vhaddad@ad.ucla.edu). Huebner: [paul.huebner.phd@anderson.ucla.edu](mailto:paul.huebner.phd@anderson.ucla.edu). Loualiche: [eloualic@umn.edu](mailto:eloualic@umn.edu).

# Contents

<b>A</b>	<b>Equilibrium Model of Financial Markets with Investor Competition</b>	<b>1</b>
A.1	The effect of a rise in passive investing . . . . .	1
A.2	Asymmetry of mispricing . . . . .	1
A.3	Limits to arbitrage . . . . .	4
A.4	The effect of a rise in passive investing under heterogeneity . . . . .	5
A.5	Exogenous change in the baseline elasticity . . . . .	7
<b>B</b>	<b>Model of Information Acquisition</b>	<b>8</b>
B.1	Setup . . . . .	8
B.2	Solving prices given information . . . . .	8
B.3	Demand elasticity . . . . .	10
B.4	Optimal information . . . . .	10
B.5	Flexibility in information acquisition . . . . .	13
B.6	Price informativeness . . . . .	14
<b>C</b>	<b>Other Foundations for the Degree of Strategic Response <math>\chi</math></b>	<b>15</b>
C.1	Learning from prices . . . . .	15
C.2	Price impact . . . . .	18
C.3	Imperfect information . . . . .	20
C.4	Partial equilibrium thinking . . . . .	20
<b>D</b>	<b>Data</b>	<b>22</b>
D.1	Institutional holdings data . . . . .	22
D.2	CRSP and Compustat . . . . .	22
D.3	Measuring passive investing . . . . .	22
D.4	Additional data definitions . . . . .	23
<b>E</b>	<b>Identification Strategy</b>	<b>25</b>
E.1	Moment conditions . . . . .	25
E.2	Solving the reflection problem . . . . .	25
E.3	Numerical procedure . . . . .	27
<b>F</b>	<b>Trading Big and Small Stocks</b>	<b>30</b>
<b>G</b>	<b>Appendix Tables</b>	<b>32</b>
<b>H</b>	<b>Appendix Figures</b>	<b>36</b>

# A Equilibrium Model of Financial Markets with Investor Competition

We derive the implications of the model of Section 2. First, we provide formal derivations for the effect of the rise of passive investing presented in Section 2.3. Then, we show how the degree of strategic response shapes the relative prevalence of overpricing and underpricing. Next, we illustrate the role of investor interactions for dynamic aspects of limits to arbitrage. Finally, we re-evaluate the effect of a rise of passive investing under heterogeneity.

## A.1 The effect of a rise in passive investing

In Section 2.3, we ask how the aggregate elasticity changes when a fraction of investors becomes passive. We start with an economy with homogeneous investors and assume an initial equilibrium  $p = \bar{p}$ . Because all investors are identical, we have  $\mathcal{E}_i = \mathcal{E}_{agg}$ . Combined with elasticity decision equation (4), this gives  $\underline{\mathcal{E}}_i = (1 + \chi)\mathcal{E}_i$ . We denote the initial equilibrium elasticity of investors by  $\mathcal{E}_i = \mathcal{E}_{agg} = \mathcal{E}_0$ .

Assume that a fraction  $1 - \alpha$  of investors becomes passive such that their elasticity becomes zero and their level of demand is unchanged. We have to determine the new elasticity of active investors  $\mathcal{E}_i$ . The individual decisions and equilibrium conditions are:

$$\mathcal{E}_i = \underline{\mathcal{E}}_i - \chi\mathcal{E}_{agg}, \tag{IA.1}$$

$$\mathcal{E}_{agg} = \alpha\mathcal{E}_i + (1 - \alpha) \cdot 0. \tag{IA.2}$$

Solving this system gives aggregate elasticity:

$$\mathcal{E}_{agg} = \frac{\alpha}{1 + \alpha\chi}\mathcal{E}_i. \tag{IA.3}$$

$$\mathcal{E}_{agg} = \frac{\alpha}{1 + \alpha\chi}(1 + \chi)\mathcal{E}_0 = \alpha\mathcal{E}_0 + \frac{\alpha\chi}{1 + \alpha\chi}(1 - \alpha)\mathcal{E}_0. \tag{IA.4}$$

The first part of the equation corresponds to the direct effect of the rise in passive investing, and the second part represents the compensation from the strategic response of the remaining active investors.

## A.2 Asymmetry of mispricing

In this section, we show that in markets with strong strategic responses, prices are less responsive to demand when their levels are high than when they are low. The opposite happens for low degrees of strategic response. This distinction has important practical implications: when strategic responses are mild, as we find in the data, we expect to observe more situations of overpricing than underpricing, and more sensitivity to demand shocks for overpriced assets. This prediction lines up with the evidence in Stambaugh, Yu, and Yuan (2012) and Stambaugh, Yu, and Yuan (2015) across a large set of anomaly strategies.<sup>1</sup> In particular, they document that trading the short leg of the portfolio (overpriced stocks)

---

<sup>1</sup>The anomaly strategies are portfolios sorted on characteristics that predict unconditional returns, in the style of Fama and French (1993).

is more profitable than the long leg (underpriced stocks) in episodes of high investment sentiment (demand shocks in our theory).<sup>2</sup>

**Multiplier.** First, we revisit the calculation of equation (3) in the presence of strategic responses. We show below (Appendix A.2.1) that the price multiplier to an aggregate demand shock becomes:

$$M_{agg} = \frac{1}{\mathcal{E}_{agg}} \cdot \frac{1}{1 + \frac{\chi}{1+\chi} \frac{\text{Var}[\mathcal{E}_i]}{\mathcal{E}_{agg}} (p - \bar{p})}, \quad (\text{IA.5})$$

where  $\text{Var}[\mathcal{E}_i]$  is the demand-weighted cross-sectional variance of elasticity.<sup>3</sup> For small deviations of the price from its baseline (small values of  $p - \bar{p}$ ), the response of prices to a change in demand is still determined by the aggregate elasticity  $\mathcal{E}_{agg}$ . For example, at first order, the effect of a rise in passive investing on aggregate elasticity is reflected one-to-one into the multiplier. In “fiercely competitive markets,” the rise in passive investing does not affect the sensitivity of prices to demand, whereas it increases it when there are no strategic responses. We now turn to how the multiplier changes when the price deviates from its baseline  $\bar{p}$ .

**No strategic response.** Without strategic responses ( $\chi = 0$ ), the aggregate multiplier is  $M_{agg} = \mathcal{E}_{agg}^{-1}$ . While each of the investors’ elasticities is fixed at  $\underline{\mathcal{E}}_i$ , their contribution to aggregate elasticity depends on their relative demand. When the price is below its baseline, the more elastic investors have a stronger response than the less elastic investors: they buy relatively more of the asset. Thus more elastic investors represent a larger fraction of the market; aggregate elasticity increases. For example, in response to a supply shock, the aggregate elasticity moves in the opposite direction from the price:  $\Delta \mathcal{E}_{agg} = -\text{Var}[\mathcal{E}_i] \Delta p$ . How much investors differ from each other, the variance of individual elasticities, controls the strength of the composition effect. The market has a higher capacity to absorb demand shocks on the downside than on the upside. Without strategic responses, overpricing is more likely to happen than underpricing.

**Strong strategic responses.** Consider now the other extreme of a high degree of strategic response,  $\chi \rightarrow \infty$ . In this case the aggregate elasticity  $\mathcal{E}_{agg}$  is constant; if more elastic investors decrease their relative share, everybody becomes more elastic and exactly compensates the initial decline. However, the multiplier changes because the second term of the product in equation (IA.5) is not constant: it is smaller than one for  $p > \bar{p}$  and larger than one otherwise. What happens then? If the price is above its baseline, the increase in individual elasticities due to the strategic response implies a decrease in demand. Because all investors become more aggressive, they demand less of the asset, which prevents the price from increasing much, leading to a smaller multiplier. The demand effect goes in the

---

<sup>2</sup>As such, imperfect competition is an alternative explanation to short-sell constraints (for example, Miller (1977) or Haddad, Ho, and Loualiche (2021)) for the pervasiveness of overpricing.

<sup>3</sup>Formally this corresponds to

$$\text{Var}[\mathcal{E}_i] = \int \left( \mathcal{E}_i^2 \frac{D_i}{S} \right) - \mathcal{E}_{agg}^2.$$

opposite direction when the price is below its baseline. Overall this leads to an opposite behavior of the multiplier  $M_{agg}$  relative to no strategic response. The multiplier is larger with underpricing than overpricing: underpricing is more likely to happen than overpricing.

### A.2.1 Derivations

How do prices in the two-layer equilibrium decrease in response to an exogenous shift in the supply of a stock,  $\Delta S$ ? We start by defining aggregation operators which weight investors' outcomes by the fraction of the stock they own:

$$\check{\mathbf{E}}[x_i] = \int x_i \frac{D_i}{S}, \quad (\text{IA.6})$$

$$\check{\mathbf{V}}\text{ar}[x_i] = \int x_i^2 \frac{D_i}{S} - \left( \int x_i \frac{D_i}{S} \right)^2 \quad (\text{IA.7})$$

$$\check{\mathbf{C}}\text{ov}[x_i, y_i] = \int (x_i - \check{\mathbf{E}}[x_i])(y_i - \check{\mathbf{E}}[y_i]) \frac{D_i}{S}. \quad (\text{IA.8})$$

To see the impact of the change in supply on the equilibrium we start by deriving the change in individual demand:

$$\Delta D_i = \Delta e^{\underline{d}_i - \mathcal{E}_i(p - \bar{p})} = -\mathcal{E}_i D_i \Delta p - (p - \bar{p}) D_i \Delta \mathcal{E}_i, \quad (\text{IA.9})$$

such that change in aggregate demand, which corresponds to the change in aggregate supply, reads:

$$\Delta S = \Delta D = -\Delta p S \int \mathcal{E}_i D_i / S - (p - \bar{p}) S \int \Delta \mathcal{E}_i D_i / S \quad (\text{IA.10})$$

$$= -\Delta p S \mathcal{E}_{agg} + \chi(p - \bar{p}) S \Delta \mathcal{E}_{agg}, \quad (\text{IA.11})$$

where in the last equation we used the individual elasticity equation  $\mathcal{E}_i = \underline{\mathcal{E}}_i - \chi \mathcal{E}_{agg}$  leading to the uniform change in individual investors' elasticity

$$\Delta \mathcal{E}_i = -\chi \Delta \mathcal{E}_{agg}. \quad (\text{IA.12})$$

To estimate the change in aggregate elasticity we write:

$$\Delta \mathcal{E}_{agg} = \int \Delta (\mathcal{E}_i D_i / S) = \int \Delta \mathcal{E}_i D_i / S - \frac{\Delta S}{S} \mathcal{E}_{agg} + \int \mathcal{E}_i \Delta D_i / S \quad (\text{IA.13})$$

$$= -\chi \Delta \mathcal{E}_{agg} - \frac{\Delta S}{S} \mathcal{E}_{agg} - \Delta p \check{\mathbf{E}}[\mathcal{E}_i^2] + \chi(p - \bar{p}) \mathcal{E}_{agg} \Delta \mathcal{E}_{agg}. \quad (\text{IA.14})$$

$$\Delta \mathcal{E}_{agg} = -\frac{1}{1 + \chi - \chi(p - \bar{p}) \mathcal{E}_{agg}} \left( \mathcal{E}_{agg} \frac{\Delta S}{S} + \check{\mathbf{E}}[\mathcal{E}_i^2] \Delta p \right), \quad (\text{IA.15})$$

Then we can plug the change in aggregate elasticity into the main equation expressing the change in aggregate demand and solve for  $\Delta p$  as a function of the change in supply  $\Delta S$ .

$$\frac{\Delta S}{S} = -\mathcal{E}_{agg} \Delta p - \frac{\chi(p - \bar{p})}{1 + \chi - \chi(p - \bar{p}) \mathcal{E}_{agg}} \left( \mathcal{E}_{agg} \frac{\Delta S}{S} + \check{\mathbf{E}}[\mathcal{E}_i^2] \Delta p \right) \quad (\text{IA.16})$$

After some algebra, we find the change in prices in response to an exogenous change in supply:

$$\Delta p = \frac{1}{\mathcal{E}_{agg}} \cdot \frac{1}{1 + \frac{\chi}{1+\chi} \frac{p-\bar{p}}{\mathcal{E}_{agg}} \check{\text{Var}}[\mathcal{E}_i]} \frac{\Delta S}{S} = M_{agg} \cdot \frac{\Delta S}{S}, \quad (\text{IA.17})$$

where we define the aggregate multiplier  $M_{agg}$  as:

$$M_{agg} = \frac{1}{\mathcal{E}_{agg}} \cdot \frac{1}{1 + \frac{\chi}{1+\chi} \frac{p-\bar{p}}{\mathcal{E}_{agg}} \check{\text{Var}}[\mathcal{E}_i]}. \quad (\text{IA.18})$$

### A.3 Limits to arbitrage

An important insight is that engaging in arbitrage trades (or more broadly exploiting mispricing) is a risky activity and this risk limits the effectiveness of arbitrageurs (De Long et al. (1990), Shleifer and Summers (1990), Shleifer and Vishny (1997), Brunnermeier and Pedersen (2008)).

If an asset is underpriced ( $p < \bar{p}$ ), we expect arbitrageurs (high elasticity investors) to take on large positions. When the mispricing worsens, the arbitrageurs suffer large losses due to their large exposure. If they are unable to raise additional capital, they have to liquidate some of their positions which pushes the price down even further. This feedback creates a natural instability of arbitrage activity: shocks that worsen the mispricing hurt the arbitrageurs and deepen the mispricing. Mitchell, Pulvino, and Stafford (2002) and Mitchell and Pulvino (2012) document this instability in action.

The degree of strategic response plays an important role in this process. Without strategic responses, the arbitrage capacity destroyed by arbitrageurs' losses is gone altogether. With strategic responses, other investors become more elastic in response to the decline of the arbitrageurs. They purchase more of the asset, thereby partially compensating the lower positions of the arbitrageurs.

We can illustrate this mechanism in our model by considering how the price responds to a demand shock that affects disproportionately investors with high elasticity. Below (Appendix A.3.1), we assume that each investor's baseline demand changes by an amount  $\{\Delta \underline{d}_i\}_i$  and compute the equilibrium price response:

$$\Delta p = M_{agg} \times \left[ \mathbf{E}(\Delta \underline{d}_i) + \frac{\chi}{1+\chi} (p - \bar{p}) \text{Cov}(\mathcal{E}_i, \Delta \underline{d}_i) \right], \quad (\text{IA.19})$$

where expectation and covariance represent demand-weighted moments. Without strategic responses,  $\chi = 0$ , the price response is simply the product of the aggregate multiplier  $M_{agg}$  with the average demand shock  $\mathbf{E}(\Delta \underline{d}_i)$ . A shock that hurts disproportionately investors with large positions (for example the asset is underpriced and these investors have a high demand elasticity) pushes the price down. The second term captures the role of strategic responses. When the asset is underpriced ( $p < \bar{p}$ ), and demand decreases more for high elasticity investors ( $\text{Cov}(\mathcal{E}_i, \Delta \underline{d}_i) < 0$ ), we obtain a positive response which compensates the direct effect. A symmetric compensation occurs when the asset is overpriced (see the derivations below). Interestingly, the stabilizing role of strategic responses is stronger the further away prices are from their baseline.<sup>4</sup>

<sup>4</sup>Duffie (2010) and Duffie and Strulovici (2012) study how the competitive response unfolds over time.

### A.3.1 Derivations

We consider an experiment where demand across investors changes by  $\{\Delta d_i\}_i$ . Solving for the equilibrium response of the price and elasticity is similar to the simple case of a change in aggregate supply. First, we consider the effect of a change in investors' demands on aggregate demand:

$$\Delta D = \Delta \int e^{d_i - \mathcal{E}_i(p - \bar{p})} = \int \Delta d_i D_i - \int \mathcal{E}_i \Delta p D_i - \int (p - \bar{p}) \Delta \mathcal{E}_i D_i \quad (\text{IA.20})$$

$$\frac{\Delta D}{S} = \check{\mathbf{E}}[\Delta d_i] - \mathcal{E}_{agg} \Delta p + \chi(p - \bar{p}) \Delta \mathcal{E}_{agg}, \quad (\text{IA.21})$$

where we have used  $\check{\mathbf{E}}[\mathcal{E}_i] = \mathcal{E}_{agg}$  and  $\Delta \mathcal{E}_i = -\chi \Delta \mathcal{E}_{agg}$ . To solve for the change in price as a function of the change in investors' demands we use the second market clearing condition of aggregate elasticities to find  $\Delta \mathcal{E}_{agg}$ :

$$\Delta \mathcal{E}_{agg} = \int \Delta \mathcal{E}_i D_i / S + \int \mathcal{E}_i \Delta D_i / S \quad (\text{IA.22})$$

$$= -\chi \Delta \mathcal{E}_{agg} + \int \mathcal{E}_i \Delta d_i D_i / S - \Delta p \int \mathcal{E}_i^2 D_i / S - (p - \bar{p}) \int \mathcal{E}_i \Delta \mathcal{E}_i D_i / S \quad (\text{IA.23})$$

$$= (-\chi + \chi(p - \bar{p}) \mathcal{E}_{agg}) \Delta \mathcal{E}_{agg} + \int \mathcal{E}_i \Delta d_i D_i / S - \Delta p \int \mathcal{E}_i^2 D_i / S \quad (\text{IA.24})$$

$$= \frac{1}{1 + \chi - \chi(p - \bar{p}) \mathcal{E}_{agg}} \cdot (\check{\mathbf{E}}[\mathcal{E}_i \Delta d_i] - \check{\mathbf{E}}[\mathcal{E}_i^2] \Delta p). \quad (\text{IA.25})$$

We plug this expression back into the expression for the change in aggregate demand above:

$$\frac{\Delta D}{S} = \check{\mathbf{E}}[\Delta d_i] - \mathcal{E}_{agg} \Delta p + \frac{\chi(p - \bar{p})}{1 + \chi - \chi p \mathcal{E}_{agg}} \cdot (\check{\mathbf{E}}[\mathcal{E}_i \Delta d_i] - \check{\mathbf{E}}[\mathcal{E}_i^2] \Delta p) \quad (\text{IA.26})$$

The supply is fixed such that  $\Delta D/S = 0$ , which after rearranging terms gives the final expression for the change in price:

$$\Delta p = \frac{1}{\mathcal{E}_{agg}} \frac{1}{1 + \frac{\chi}{1 + \chi} \frac{p - \bar{p}}{\mathcal{E}_{agg}} \text{Var}[\mathcal{E}_i]} \left( \check{\mathbf{E}}[\Delta d_i] + \frac{\chi}{1 + \chi} (p - \bar{p}) \check{\text{Cov}}(\mathcal{E}_i, \Delta d_i) \right) \quad (\text{IA.27})$$

We recognize the first term as the standard aggregate multiplier (obtained when we derived the response to change in supply) and write the price response as:

$$\Delta p = M_{agg} \cdot \left( \underbrace{\check{\mathbf{E}}[\Delta d_i]}_{\text{average demand shock}} + \frac{\chi}{1 + \chi} (p - \bar{p}) \underbrace{\check{\text{Cov}}(\mathcal{E}_i, \Delta d_i)}_{\text{average elasticity composition}} \right). \quad (\text{IA.28})$$

## A.4 The effect of a rise in passive investing under heterogeneity

In Section A.1, we show how the elasticity changes when a fraction of investors become passive. Here, we revisit this calculation under heterogeneity in baseline elasticity  $\mathcal{E}_i$  (see

also Section 5.1) and degree of strategic response  $\chi_i$ . Again, we start with an economy with an initial equilibrium  $p = \bar{p}$ .

Individual decisions and equilibrium conditions are:

$$\mathcal{E}_i = \underline{\mathcal{E}}_i - \chi_i \mathcal{E}_{agg}, \quad (\text{IA.29})$$

$$\mathcal{E}_{agg} = \int \mathcal{E}_i \frac{D_i}{S}. \quad (\text{IA.30})$$

Combining the two yields

$$\mathcal{E}_{agg} = \int \mathcal{E}_i \frac{D_i}{S} \quad (\text{IA.31})$$

$$= \int \underline{\mathcal{E}}_i \frac{D_i}{S} - \mathcal{E}_{agg} \int \chi_i \frac{D_i}{S} \quad (\text{IA.32})$$

$$= \check{\mathbf{E}}[\underline{\mathcal{E}}_i] - \check{\mathbf{E}}[\chi_i] \mathcal{E}_{agg} \quad (\text{IA.33})$$

$$= \check{\mathbf{E}}[\underline{\mathcal{E}}_i] \times \frac{1}{1 + \check{\mathbf{E}}[\chi_i]}. \quad (\text{IA.34})$$

Denote the set of active investors as *Active*, and the fraction of assets held by this group by  $|Active|$ . Passive investors have  $\underline{\mathcal{E}}_i = \chi_i = 0$ , such that

$$\mathcal{E}_{agg} = |Active| \times \check{\mathbf{E}}[\underline{\mathcal{E}}_i | i \in Active] \times \frac{1}{1 + |Active| \times \check{\mathbf{E}}[\chi_i | i \in Active]}. \quad (\text{IA.35})$$

The aggregate elasticity is the product of the active share, the average baseline elasticity among active investors, and an equilibrium term capturing the strategic responses of active investors. Note that this expression looks exactly like in the case with homogeneous  $\chi$  (e.g. equation (25)), other than now the strength of strategic response is controlled by the average  $\chi_i$  among active investors.

Now consider the effect of a change in the active share  $|Active|$ . We assume that random investors are switching, meaning that the set of active investors that become passive is a representative sample of the active population, such that  $\check{\mathbf{E}}[\underline{\mathcal{E}}_i | i \in Active]$  and  $\check{\mathbf{E}}[\chi_i | i \in Active]$  remain unaffected by the change. Then:

$$\frac{d \log \mathcal{E}_{agg}}{d \log |Active|} = \frac{|Active|}{\mathcal{E}_{agg}} \times \left[ \frac{\check{\mathbf{E}}[\underline{\mathcal{E}}_i | i \in Active] - \frac{|Active| \times \check{\mathbf{E}}[\underline{\mathcal{E}}_i | i \in Active] \times \check{\mathbf{E}}[\chi_i | i \in Active]}{1 + |Active| \times \check{\mathbf{E}}[\chi_i | i \in Active]}}{1 + |Active| \times \check{\mathbf{E}}[\chi_i | i \in Active]} \right] \quad (\text{IA.36})$$

$$= 1 - \frac{|Active| \times \check{\mathbf{E}}[\chi_i | i \in Active]}{1 + |Active| \times \check{\mathbf{E}}[\chi_i | i \in Active]} \quad (\text{IA.37})$$

$$= \frac{1}{1 + |Active| \times \check{\mathbf{E}}[\chi_i | i \in Active]}. \quad (\text{IA.38})$$

The pass-through from a change in the active share mirrors equation (26). The only difference is that the average degree of strategic response  $\chi_i$  among active investors is what determines the pass-through instead of the single parameter  $\chi$  of the homogenous model.



## A.5 Exogenous change in the baseline elasticity

Last, we consider a change in investors' baseline elasticity  $\Delta\underline{\mathcal{E}}_i$ . We start with the relative change in aggregate demand:

$$\frac{\Delta D}{S} = -(p - \bar{p}) \int \Delta\underline{\mathcal{E}}_i D_i/S - \mathcal{E}_{agg} \Delta p \quad (\text{IA.39})$$

$$= -(p - \bar{p}) \check{\mathbf{E}}[\Delta\underline{\mathcal{E}}_i] + \chi(p - \bar{p}) \Delta\mathcal{E}_{agg} - \mathcal{E}_{agg} \Delta p \quad (\text{IA.40})$$

We use market clearing to get  $\Delta D/S = 0$ :

$$\Delta p = -\frac{1}{\mathcal{E}_{agg}} \left( (p - \bar{p}) \check{\mathbf{E}}[\Delta\underline{\mathcal{E}}_i] - \chi(p - \bar{p}) \Delta\mathcal{E}_{agg} \right). \quad (\text{IA.41})$$

The change in aggregate elasticity in the case of changes in baseline elasticities is different than in the previous cases:

$$\Delta\mathcal{E}_{agg} = \int \Delta(\mathcal{E}_i D_i/S) = \int \Delta\underline{\mathcal{E}}_i D_i/S - \chi \Delta\mathcal{E}_{agg} + \int \mathcal{E}_i \Delta D_i/S \quad (\text{IA.42})$$

$$= \int \Delta\underline{\mathcal{E}}_i D_i/S - \chi \Delta\mathcal{E}_{agg} - (p - \bar{p}) \int \mathcal{E}_i \Delta\underline{\mathcal{E}}_i D_i/S + \chi(p - \bar{p}) \mathcal{E}_{agg} \Delta\mathcal{E}_{agg} - \Delta p \int \mathcal{E}_i^2 D_i/S \quad (\text{IA.43})$$

$$= \check{\mathbf{E}}[\Delta\underline{\mathcal{E}}_i] - \chi \Delta\mathcal{E}_{agg} - (p - \bar{p}) \check{\mathbf{E}}[\mathcal{E}_i \Delta\underline{\mathcal{E}}_i] + \chi(p - \bar{p}) \mathcal{E}_{agg} \Delta\mathcal{E}_{agg} - \Delta p \check{\mathbf{E}}[\mathcal{E}_i^2] \quad (\text{IA.44})$$

After rearranging the terms and plugging in the expression for the change in the price, we find:

$$\left( 1 + \chi + \chi(p - \bar{p}) \frac{\check{\text{Var}}[\mathcal{E}_i]}{\mathcal{E}_{agg}} \right) \Delta\mathcal{E}_{agg} = \left( 1 + (p - \bar{p}) \frac{\check{\text{Var}}[\mathcal{E}_i]}{\mathcal{E}_{agg}} \right) \check{\mathbf{E}}[\Delta\underline{\mathcal{E}}_i] - (p - \bar{p}) \check{\text{Cov}}(\mathcal{E}_i, \Delta\underline{\mathcal{E}}_i) \quad (\text{IA.45})$$

Plugging back into the change in the price we have:

$$\Delta p = -M_{agg} \frac{p - \bar{p}}{1 + \chi} \left( \check{\mathbf{E}}[\Delta\underline{\mathcal{E}}_i] + (p - \bar{p}) \check{\text{Cov}}(\mathcal{E}_i, \Delta\underline{\mathcal{E}}_i) \right). \quad (\text{IA.46})$$

## B Model of Information Acquisition

### B.1 Setup

There is one period and one asset, and a continuum of agents indexed by  $i \in [0, 1]$ . Each agent has CARA preferences with risk aversion  $\rho_i$ :

$$U_i = \mathbf{E}_i[-e^{-\rho_i W_i}], \quad (\text{IA.47})$$

and initial wealth  $W_i$ . The gross risk-free rate is 1, and the random asset payoff is  $f$ . The asset is in noisy supply  $\bar{x} + x$  with  $\bar{x}$  an exogenous fixed parameter and  $x \sim \mathcal{N}(0, \sigma_x^2)$ .

Each agent has a prior that  $f \sim \mathcal{N}(\mu_i, \sigma_i^2)$ . Following Veldkamp (2011), agents start with a flat prior on  $f$  and receive signal  $\mu_i$  such that the signal is distributed  $\mu_i \sim \mathcal{N}(f, \sigma_i^2)$ . Each agent can acquire a private signal  $\eta_i \sim \mathcal{N}(f, \sigma_{i,\eta}^2)$  at cost  $c_i(\sigma_i^{-2} + \sigma_{i,\eta}^{-2})$ , with  $c_i(\cdot)$  a non-decreasing positive function. That is, obtaining more precise signals is more costly. The signal being private implies in particular that signal realizations are uncorrelated across agents conditional on the fundamental  $f$ .

We focus on rational expectations equilibria, and among those linear equilibria specifically. These are equilibria in which the price takes the form:

$$p = A + Bf + Cx. \quad (\text{IA.48})$$

An equilibrium is a set of coefficient  $(A, B, C)$ , information choices  $\sigma_{i,\eta}^2$ , demand curves  $D_i(p|\eta_i)$  such that:

- (a) Each demand function and information choice maximizes expected utility, taking as given the price function.
- (b) The market for the asset clears:  $\bar{x} + x = \int D_i(p|\eta_i) di$ .

To solve the model, we proceed in two steps: first we solve for the price given information decisions; second, we derive equilibrium information decisions.

### B.2 Solving prices given information

We are going to solve for the price function  $p = A + Bf + Cx$ . First, we solve for allocations given the information choice and finally we use market clearing to pin down the price.

Agents form posterior beliefs on the fundamental  $f$  based on their prior  $\mu_i$ , signal  $\eta_i$ , and based on prices. The signal agents can extract from prices about  $f$  is:

$$s(p) = \frac{p - A}{B} = f + \frac{C}{B}x. \quad (\text{IA.49})$$

Given the three signals, we are able to derive the posterior belief about  $f$ , which will be distributed as  $\mathcal{N}(\hat{\mu}_i, \hat{\sigma}_i^2)$  as follows:

$$\hat{\sigma}_i^{-2} = \sigma_i^{-2} + \sigma_{i\eta}^{-2} + \frac{B^2}{C^2} \sigma_x^{-2} \quad (\text{IA.50})$$

$$\hat{\mu}_i = \hat{\sigma}_i^2 \left( \sigma_i^{-2} \mu_i + \sigma_{i\eta}^{-2} \eta_i + \frac{B^2}{C^2} \sigma_x^{-2} s(p) \right) \quad (\text{IA.51})$$

**Asset Demand.** Abstracting from the cost of acquiring information, the expected utility function for a given asset holding  $q_i$  is:

$$U_i(q_i) = -\mathbf{E}[\exp(-\rho_i(fq_i - pq_i))] \quad (\text{IA.52})$$

$$= -\exp\left(-\rho_i q_i (\mathbf{E}[f] - p) + \frac{\rho_i^2}{2} q_i^2 \text{Var}[f]\right). \quad (\text{IA.53})$$

The first order condition with respect to  $q_i$  gives us immediately:

$$\begin{aligned} & -\rho_i (\mathbf{E}[f] - p) + \rho_i^2 q_i \text{Var}[f] = 0 \\ \iff & q_i = \frac{1}{\rho_i \text{Var}[f]} (\mathbf{E}[f] - p) \\ \iff & q_i = \frac{1}{\rho_i} \hat{\sigma}_i^{-2} (\hat{\mu}_i - p) \end{aligned} \quad (\text{IA.54})$$

**Market Clearing.** The market clearing condition reads:

$$\int q_i di = \bar{x} + x. \quad (\text{IA.55})$$

Given asset demand this translates into:

$$\int \frac{1}{\rho_i} \hat{\sigma}_i^{-2} (\hat{\mu}_i - p) di = \bar{x} + x \quad (\text{IA.56})$$

The goal now is to find  $(A, B, C)$ , which we identify directly from the market clearing condition. First we replace the expressions for the price function and the posteriors mean and variances in the market clearing equation:

$$\int \frac{1}{\rho_i} \hat{\sigma}_i^{-2} \left[ f + \frac{B}{C} \frac{\hat{\sigma}_i^2}{\sigma_x^2} x \right] di - \int \frac{1}{\rho_i} \hat{\sigma}_i^{-2} [A + Bf + Cx] di = \bar{x} + x \quad (\text{IA.57})$$

We identify all the terms that are linear in  $f$  and find that  $B = 1$ . Next, we group the terms that are linear in  $x$ , which yields

$$\begin{aligned} & \int \frac{1}{\rho_i} \frac{1}{C} \sigma_x^{-2} di - \int \frac{1}{\rho_i} \hat{\sigma}_i^{-2} C di = 1 \\ \iff & \int \frac{1}{\rho_i} \left[ \frac{1}{C} \sigma_x^{-2} - C \sigma_i^{-2} - C \sigma_{in}^{-2} - \frac{1}{C} \sigma_x^{-2} \right] di = 1 \\ \iff & C = - \left[ \int \frac{1}{\rho_i} (\sigma_i^{-2} + \sigma_{in}^{-2}) di \right]^{-1}, \end{aligned} \quad (\text{IA.58})$$

where we used the expression of the posterior found above to substitute into the second equation. Last, we gather the constant terms to find  $A$ :

$$A = -\bar{x} \left[ \int \frac{1}{\rho_i} \hat{\sigma}_i^{-2} di \right]^{-1}. \quad (\text{IA.59})$$

### B.3 Demand elasticity

We recall the demand schedule for agent  $i$ :

$$\begin{aligned}
q_i &= \frac{1}{\rho_i} \hat{\sigma}_i^{-2} (\hat{\mu}_i - p) \\
&= \frac{1}{\rho_i} \hat{\sigma}_i^{-2} (\hat{\sigma}_i^2 [\sigma_i^{-2} \mu_i + \sigma_{i\eta}^{-2} \eta_i + C^{-2} \sigma_x^{-2} s(p)] - p) \\
&= \frac{1}{\rho_i} (\sigma_i^{-2} \mu_i + \sigma_{i\eta}^{-2} \eta_i + C^{-2} \sigma_x^{-2} (p - A) - \hat{\sigma}_i^{-2} p) \\
&= \frac{1}{\rho_i} (\sigma_i^{-2} \mu_i + \sigma_{i\eta}^{-2} \eta_i + (C^{-2} \sigma_x^{-2} - \hat{\sigma}_i^{-2}) p - C^{-2} \sigma_x^{-2} A). \tag{IA.60}
\end{aligned}$$

We can read the demand elasticity off the terms proportional to the price  $p$  as:

$$\mathcal{E}_i = -\frac{dq_i}{dp} = -\frac{1}{\rho_i} (C^{-2} \sigma_x^{-2} - \hat{\sigma}_i^{-2}) = \frac{1}{\rho_i} (\sigma_i^{-2} + \sigma_{i\eta}^{-2}). \tag{IA.61}$$

In the model, a regression of  $q_i$  on  $p$  would not give us the proper elasticity. There is a bias in the regression because  $p$  is correlated with  $\mu_i$  and  $\eta_i$ . It is still possible to recover the elasticity using an instrument; for example the supply shock  $x$  covaries with  $p$  but is uncorrelated with  $\mu_i$  and  $\eta_i$ .

We define the aggregate demand elasticity as

$$\mathcal{E}_{agg} = \int \mathcal{E}_j dj. \tag{IA.62}$$

We can express the equilibrium in terms of demand elasticities. Taking (IA.61) and (IA.62) together, we express the solution for the equilibrium

$$C = - \left[ \int \frac{1}{\rho_j} (\sigma_j^{-2} + \sigma_{j\eta}^{-2}) dj \right]^{-1} = \mathcal{E}_{agg}^{-1}. \tag{IA.63}$$

### B.4 Optimal information

**Computing expected utility.** Conditional on the signal and the price, expected utility is:

$$U_i(q_i) = -\mathbf{E} [\exp(-\rho_i (fq_i - pq_i)) | p, \eta] \tag{IA.64}$$

$$= -\exp \left( -\rho_i q_i (\mathbf{E}[f|p, \eta] - p) + \frac{\rho_i^2}{2} q_i^2 \text{Var}[f|p, \eta] \right) \tag{IA.65}$$

$$= -\exp \left( -\frac{1}{2} \frac{(\mathbf{E}[f|p, \eta] - p)^2}{\text{Var}[f|p, \eta]} \right), \tag{IA.66}$$

where the last line is derived using standard properties of quadratic functions.<sup>5</sup>

---

<sup>5</sup>For a function  $f(x) = ax^2 + bx$ , the maximum is reached for  $x^* = -b/(2a)$  and its value is  $f(x^*) = -b^2/(4a)$ .

We can write:

$$\mathbf{E}[f|p, \eta] - p = \underbrace{(\mathbf{E}[f|p, \eta] - \mathbf{E}[f|p])}_z + (\mathbf{E}[f|p] - p). \quad (\text{IA.67})$$

Conditional on  $p$ ,  $z$  has mean 0 and its variance  $\sigma_z^2$  can be obtained from:

$$\underbrace{f - \mathbf{E}[f|p]}_{\text{variance: } (\sigma_i^{-2} + \sigma_x^{-2}/C^2)^{-1}} = \underbrace{(f - \mathbf{E}[f|p, \eta])}_{\text{variance: } \hat{\sigma}_i^2} + \underbrace{z}_{\text{variance: } \sigma_z^2} \quad (\text{IA.68})$$

Using equation (7.32) in Veldkamp (2011), this maps to:<sup>6</sup>

$$F = -\frac{1}{2} \frac{1}{\hat{\sigma}_i^2} \quad (\text{IA.69})$$

$$G = -(\mathbf{E}[f|p] - p) \frac{1}{\hat{\sigma}_i^2} \quad (\text{IA.70})$$

$$H = -\frac{1}{2} (\mathbf{E}[f|p] - p)^2 \frac{1}{\hat{\sigma}_i^2} \quad (\text{IA.71})$$

So expected utility conditional on the price is:

$$\begin{aligned} U_0|_p &= -(1 - 2\sigma_z^2 F)^{-1/2} \exp\left(\frac{1}{2} G^2 (1 - 2\sigma_z^2 F)^{-1} \sigma_z^2 + H\right) \\ &= -\left(1 + \frac{\sigma_z^2}{\hat{\sigma}_i^2}\right)^{-1/2} \exp\left(\frac{1}{2} \frac{(\mathbf{E}[f|p] - p)^2}{\hat{\sigma}_i^2} \left[\left(1 + \frac{\sigma_z^2}{\hat{\sigma}_i^2}\right)^{-1} \frac{1}{\hat{\sigma}_i^2} \sigma_z^2 - 1\right]\right) \\ &= -\left(1 + \frac{\sigma_z^2}{\hat{\sigma}_i^2}\right)^{-1/2} \exp\left(\frac{1}{2} \frac{(\mathbf{E}[f|p] - p)^2}{\hat{\sigma}_i^2} \left[\left(1 + \frac{\sigma_z^2}{\hat{\sigma}_i^2}\right)^{-1} \left(\frac{1}{\hat{\sigma}_i^2} \sigma_z^2 - 1 - \frac{\sigma_z^2}{\hat{\sigma}_i^2}\right)\right]\right) \\ &= -\left(1 + \frac{\sigma_z^2}{\hat{\sigma}_i^2}\right)^{-1/2} \exp\left(-\frac{1}{2} \frac{(\mathbf{E}[f|p] - p)^2}{\hat{\sigma}_i^2} \left[\left(1 + \frac{\sigma_z^2}{\hat{\sigma}_i^2}\right)^{-1}\right]\right) \\ U_0|_p &= -\left(1 + \frac{\sigma_z^2}{\hat{\sigma}_i^2}\right)^{-1/2} \exp\left(-\frac{1}{2} \frac{(\mathbf{E}[f|p] - p)^2}{\hat{\sigma}_i^2 + \sigma_z^2}\right). \end{aligned} \quad (\text{IA.72})$$

Expected utility is:

$$\begin{aligned} \mathbf{E}[U_0|_p] &= -\left(1 + \frac{\sigma_z^2}{\hat{\sigma}_i^2}\right)^{-1/2} \mathbf{E} \exp\left(-\frac{1}{2} \frac{(\mathbf{E}[f|p] - p)^2}{\hat{\sigma}_i^2 + \sigma_z^2}\right) \\ &= -\left[\frac{\sigma_i^{-2} + \sigma_x^{-2}/C^2}{\sigma_i^{-2} + \sigma_{in}^{-2} + \sigma_x^{-2}/C^2}\right]^{1/2} \cdot \mathbf{E} \left[\exp\left(-\frac{1}{2} \frac{(\mathbf{E}[f|p] - p)^2}{(\sigma_i^{-2} + \sigma_x^{-2}/C^2)^{-1}}\right)\right] \end{aligned} \quad (\text{IA.73})$$

where we use (IA.68) in the second equality.

<sup>6</sup>There is a general formula for the mean of the exponential of the quadratic of a normal variable. If we take the multivariate normal  $z \sim \mathcal{N}(0, \Sigma)$ :

$$\mathbf{E}[\exp(z' F z + G' z + H)] = |I - 2\Sigma F|^{-1/2} \exp\left(\frac{1}{2} G'(I - 2\Sigma F)^{-1} \Sigma G + H\right)$$

**Optimal information.** To derive the optimal information choice, investors trade off utility with the cost of acquiring information, which translates in utility terms to:

$$U_0^{(c)} = U_0 \cdot \exp(\rho_i c_i(\sigma_i^{-2} + \sigma_{i\eta}^{-2})). \quad (\text{IA.74})$$

The cost function  $c_i(\cdot)$  is increasing in the signal precision and can be specific to investor  $i$ . We obtain the first-order condition that determines the information choice:

$$\begin{aligned} & \max_{\sigma_{i\eta}^{-2}} -\log(-U_0) - \rho_i c_i(\sigma_i^{-2} + \sigma_{i\eta}^{-2}) \\ \iff & \max_{\sigma_{i\eta}^{-2}} -\log(-U_0) - \rho_i c_i(\rho_i \mathcal{E}_i) \\ \iff & \frac{1}{2} \frac{1}{\sigma_i^{-2} + \sigma_{i\eta}^{-2} + \sigma_x^{-2}/C^2} = \rho_i c'_i(\sigma_i^{-2} + \sigma_{i\eta}^{-2}). \\ \iff & \frac{1}{2} \frac{1}{\rho_i \mathcal{E}_i + \mathcal{E}_{agg}^2 \sigma_x^{-2}} = \rho_i c'_i(\rho_i \mathcal{E}_i). \end{aligned} \quad (\text{IA.75})$$

**Example 1: linear cost function.** We consider the case of a constant marginal cost for any information acquired past the initial endowment:  $c_i(x) = c_{1,i} \max(x - \sigma_i^{-2}, 0)$ . Note that in this case not all agents acquire information since  $\sigma_{i\eta}^{-2} > 0$ , so the actual precision is

$$\sigma_{i\eta}^{-2} = \max\left(\frac{1}{2\rho_i c_{1,i}} - \sigma_i^{-2} - \mathcal{E}_{agg}^2 \sigma_x^{-2}, 0\right). \quad (\text{IA.76})$$

We can rewrite the choice of information as a choice of elasticity:

$$\mathcal{E}_i = \underbrace{\frac{1}{2\rho_i^2 c_{1,i}}}_{\text{Investor characteristics}} - \frac{\sigma_x^{-2}}{\rho_i} \underbrace{\mathcal{E}_{agg}^2}_{\text{market elasticity}} \quad (\text{IA.77})$$

**Example 2: linear response to aggregate elasticity.** To relate to the model of Section 2, we ask if there is a reasonable family of cost functions that exactly gives rise to equation (4). We are looking for a cost function such that  $\mathcal{E}_i = \alpha - \beta \mathcal{E}_{agg}$ . Equivalently, this corresponds to  $\mathcal{E}_{agg} = \frac{1}{\beta}(\alpha - \mathcal{E}_i)$ . Plugging in the first order condition, this gives:

$$2\rho_i^2 c'_i(\rho_i \mathcal{E}_i) = \frac{1}{\mathcal{E}_i + \frac{\sigma_x^{-2}}{\rho_i \beta^2} (\alpha - \mathcal{E}_i)^2} \quad (\text{IA.78})$$

$$\stackrel{def}{=} \tilde{c}'_i(\mathcal{E}_i) = \frac{1}{\frac{\sigma_x^{-2}}{\rho_i \beta^2} \mathcal{E}_i^2 + \left(1 - 2\frac{\alpha \sigma_x^{-2}}{\rho_i \beta^2}\right) \mathcal{E}_i + \frac{\alpha^2 \sigma_x^{-2}}{\rho_i \beta^2}} \quad (\text{IA.79})$$

The denominator of the right-hand-side is a second degree polynomial, we solve for its roots. The discriminant is:

$$\Delta = \left(1 - 2\frac{\alpha \sigma_x^{-2}}{\rho_i \beta^2}\right)^2 - 4\frac{\sigma_x^{-2}}{\rho_i \beta^2} \frac{\alpha^2 \sigma_x^{-2}}{\rho_i \beta^2} \quad (\text{IA.80})$$

$$= 1 - 4\frac{\alpha \sigma_x^{-2}}{\rho_i \beta^2} \quad (\text{IA.81})$$

Let us assume  $\Delta < 0$ . This is equivalent to  $\rho_i\beta^2 < 4\alpha\sigma_x^{-2}$ . In this case, we have, using standard results on the primitive of the inverse of a polynomial:

$$\tilde{c}_i(\mathcal{E}_i) = \frac{2 \arctan \left( \frac{2 \frac{\sigma_x^{-2}}{\rho_i\beta^2} \mathcal{E}_i + \left(1 - 2 \frac{\alpha\sigma_x^{-2}}{\rho_i\beta^2}\right)}{\sqrt{4 \frac{\alpha\sigma_x^{-2}}{\rho_i\beta^2} - 1}} \right)}{\sqrt{4 \frac{\alpha\sigma_x^{-2}}{\rho_i\beta^2} - 1}} + K \quad (\text{IA.82})$$

The cost function is convex as long as the argument of the arctangent is negative, so:

$$\mathcal{E}_i \leq \alpha - \frac{\rho_i\beta^2}{2\sigma_x^{-2}}. \quad (\text{IA.83})$$

We can see that if the right-hand-side of this condition is positive, the condition of  $\Delta < 0$  is automatically satisfied.

After rescaling,  $2\rho_i c_i(\rho_i \mathcal{E}_i) = \tilde{c}_i(\mathcal{E}_i)$ , or equivalently  $c_i(x) = \frac{1}{2\rho_i} \tilde{c}_i(x/\rho_i)$  we have:

$$c_i(x) = \frac{1}{\rho_i} \frac{1}{\sqrt{2\alpha\tilde{\beta} - 1}} \arctan \left( \frac{\tilde{\beta} \frac{x}{\rho_i} + (1 - \alpha\tilde{\beta})}{\sqrt{2\alpha\tilde{\beta} - 1}} \right) + \tilde{K} \quad (\text{IA.84})$$

with  $\tilde{\beta} = 2\sigma_x^{-2}/(\rho_i\beta^2)$ , and the condition  $\alpha - \rho_i\beta^2/(2\sigma_x^{-2}) \geq 0$  becomes  $\alpha\tilde{\beta} \geq 1$ . We can collect these results in a proposition.

**Proposition 1.** *For any  $a > 0$  and  $b > 0$  so that  $ab > 1$ , assume the information cost follows the function:*

$$\begin{aligned} c_i(x) &= 0, \text{ if } x < 0, \\ c_i(x) &= \frac{1}{\rho_i} \frac{1}{\sqrt{2ab - 1}} \arctan \left( \frac{b \frac{x}{\rho_i} + (1 - ab)}{\sqrt{2ab - 1}} \right) + K, \text{ if } 0 \leq x/\rho_i \leq a - b^{-1} \\ c_i(x) &= +\infty, \text{ if } x/\rho_i \geq a - b^{-1}, \end{aligned} \quad (\text{IA.85})$$

where  $K$  is such that  $c_i(0) = 0$ . This cost function is increasing and convex. Then the optimal elasticity is:

$$\mathcal{E}_i = \mathcal{E}_{0,i} - \chi \mathcal{E}_{agg}, \quad (\text{IA.86})$$

with  $\mathcal{E}_{0,i} = a$  and  $\chi = \sqrt{(2\sigma_x^{-2})/(\rho_i b)}$ .

## B.5 Flexibility in information acquisition

We turn to the role of the flexibility in the acquisition of information for the degree of strategic response. We first take the inverse of (IA.75)

$$\rho_i \mathcal{E}_i + \mathcal{E}_{agg}^2 \sigma_x^{-2} = \frac{1}{2\rho_i} \cdot \frac{1}{c'_i(\rho_i \mathcal{E}_i)}. \quad (\text{IA.87})$$

Then we use the implicit function theorem:

$$\rho_i \frac{d\mathcal{E}_i}{d\mathcal{E}_{agg}} + 2\mathcal{E}_{agg}\sigma_x^{-2} = -\frac{1}{2\rho_i} \cdot \frac{\rho_i c_i''}{c_i'^2} \frac{d\mathcal{E}_i}{d\mathcal{E}_{agg}}. \quad (\text{IA.88})$$

This yields

$$\chi = -\frac{d\mathcal{E}_i}{d\mathcal{E}_{agg}} = 2\mathcal{E}_{agg}\sigma_x^{-2} \cdot \frac{1}{\rho_i + \frac{1}{2} \frac{c_i''}{c_i'^2}} \quad (\text{IA.89})$$

The response depends on the curvature of the information acquisition cost function. If the curvature is zero (as is the case in our linear cost example), then the response is highest. A larger curvature would elicit a weaker response.

## B.6 Price informativeness

We define price informativeness for investor  $i$  as the ratio of the precision of their belief about the fundamental when they condition on their private information and on the price and the precision of their belief using private information only:

$$\begin{aligned} \mathcal{I}_i &= \frac{\text{Var}(f|\mu_i, \eta_i, p)^{-1}}{\text{Var}(f|\mu_i, \eta_i)^{-1}} = \frac{\sigma_i^{-2} + \sigma_{i\eta}^{-2} + \mathcal{E}_{agg}^2 \sigma_x^{-2}}{\sigma_i^{-2} + \sigma_{i\eta}^{-2}} \\ &= 1 + \mathcal{E}_{agg} \frac{\mathcal{E}_{agg}}{\rho_i \mathcal{E}_i} \sigma_x^{-2} \end{aligned} \quad (\text{IA.90})$$

We also define the absolute price informativeness of the price as

$$\mathcal{I}_{abs} = \text{Var}(f|p)^{-1} = \mathcal{E}_{agg}^2 \sigma_x^{-2}. \quad (\text{IA.91})$$



# C Other Foundations for the Degree of Strategic Response $\chi$

## C.1 Learning from prices

We consider a model in which agents can learn from prices, which highlights a distinct mechanism from that of the previous section. Two main assumptions differ: agents cannot acquire information, and there is residual uncertainty about the asset payoff that cannot be learned. This setting leads to a new determinant of demand elasticity, beyond risk aversion and prior information. When many traders are aggressive, prices are more informative. How should one react? On the one hand, the extra information implies that price variation are less indicative of future return, and that pushes the investor to trade less aggressively. On the other hand, the extra information implies that returns appear less risky, and that pushes the investor to trade more aggressively. Increased price informativeness reveals relatively more about the fundamental than the payoff risk, exactly because of the presence of residual uncertainty.<sup>7</sup> Therefore the first effect dominates: the investor responds by being less aggressive,  $\chi > 0$ . This response is stronger when residual uncertainty is higher.

### C.1.1 Setup

The asset trades at endogenous price  $p$  and pays off  $f + \epsilon$ , with  $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$ . There is a continuum of mass 1 of agents indexed by  $i$ . Each agent has CARA preferences with risk aversion  $\rho_i$ . Each agent has a flat prior on  $f$  and receives an independent signal  $\mu_i$ , such that  $\mu_i \sim \mathcal{N}(f, \sigma_i^2)$ . The asset is in noisy supply  $\bar{x} + x$  with  $\bar{x}$  a constant and  $x \sim \mathcal{N}(0, \sigma_x^2)$ .

We look for a rational expectations equilibrium, with:

$$p = A + Bf + Cx. \tag{IA.92}$$

### C.1.2 Equilibrium

**Learning from the price.** After observing the price, agent  $i$ 's posterior belief about the fundamental  $f$  is  $\mathcal{N}(\hat{\mu}_i, \hat{\sigma}_i^2)$ , with:

$$\hat{\sigma}_i^{-2} = \sigma_i^{-2} + \frac{B^2}{C^2} \sigma_x^{-2}, \tag{IA.93}$$

$$\hat{\mu}_i = \hat{\sigma}_i^2 \left( \sigma_i^{-2} \mu_i + \frac{B^2}{C^2} \sigma_x^{-2} s(p) \right), \tag{IA.94}$$

where the signal from the price is:

$$s(p) = \frac{p - A}{B} = f + \frac{C}{B}x. \tag{IA.95}$$

---

<sup>7</sup>In the model of the previous section, the two effects exactly cancelled out. The response was coming from changes in information acquired, which is shut down here.

Taking the average over agents of type  $i$  (we use a law of large numbers in the population), we have

$$\mathbf{E}_i [\hat{\mu}_i] = \hat{\sigma}_i^2 \left( \sigma_i^{-2} f + \frac{B^2}{C^2} \sigma_x^{-2} \left[ f + \frac{C}{B} x \right] \right) \quad (\text{IA.96})$$

$$= f + \frac{B}{C} \frac{\hat{\sigma}_i^2}{\sigma_x^2} x. \quad (\text{IA.97})$$

**Asset demand.** Asset demand  $q_i$  is given by the standard optimum portfolio choice:

$$q_i = \frac{1}{\rho_i} \frac{\mathbf{E}[f + \epsilon | \mu_i, p] - p}{\text{Var}[f + \epsilon | \mu_i, p]} \quad (\text{IA.98})$$

$$= \frac{1}{\rho_i} \frac{\hat{\mu}_i - p}{\hat{\sigma}_i^2 + \sigma_\epsilon^2}. \quad (\text{IA.99})$$

**Market clearing.** The total demand for the asset must equal its supply:

$$\int q_i di = \bar{x} + x, \quad (\text{IA.100})$$

$$\int \frac{1}{\rho_i} \frac{1}{\hat{\sigma}_i^2 + \sigma_\epsilon^2} \left[ f + \frac{B}{C} \frac{\hat{\sigma}_i^2}{\sigma_x^2} x - A - Bf - Cx \right] di = \bar{x} + x. \quad (\text{IA.101})$$

This gives:

$$B = 1, \quad (\text{terms in } f) \quad (\text{IA.102})$$

$$\int \frac{1}{\rho_i} \frac{1}{\hat{\sigma}_i^2 + \sigma_\epsilon^2} \left[ \frac{B}{C} \frac{\hat{\sigma}_i^2}{\sigma_x^2} - C \right] di = 1. \quad (\text{terms in } x) \quad (\text{IA.103})$$

Plugging in the definition of  $\hat{\sigma}_i^2$ , we obtain

$$\int \frac{1}{\rho_i} \frac{1}{\hat{\sigma}_i^2 + \sigma_\epsilon^2} \left[ \frac{1}{C^2} \frac{\hat{\sigma}_i^2}{\sigma_x^2} - 1 \right] di = C^{-1}, \quad (\text{IA.104})$$

$$\int \frac{1}{\rho_i} \frac{1}{\hat{\sigma}_i^2 + \sigma_\epsilon^2} \left[ \frac{1}{C^2} \sigma_x^{-2} \frac{1}{\sigma_i^{-2} + \frac{1}{C^2} \sigma_x^{-2}} - 1 \right] di = C^{-1}, \quad (\text{IA.105})$$

$$\int \frac{1}{\rho_i} \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \sigma_\epsilon^2} \left[ \frac{1}{C^2} \sigma_x^{-2} - \sigma_i^{-2} - \frac{1}{C^2} \sigma_x^{-2} \right] di = C^{-1}. \quad (\text{IA.106})$$

Therefore we have:

$$C^{-1} = - \int \frac{1}{\rho_i} \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \sigma_\epsilon^2} \frac{1}{\sigma_i^2} di, \quad (\text{IA.107})$$

$$-C^{-1} = \int \frac{1}{\rho_i} \frac{1}{1 + \sigma_\epsilon^2 \left( \sigma_i^{-2} + \frac{1}{C^2} \sigma_x^{-2} \right)} \frac{1}{\sigma_i^2} di. \quad (\text{IA.108})$$

Define  $\tilde{C} = -C$ , which is positive. We can rewrite:

$$\tilde{C}^{-1} = \int \frac{1}{\rho_i} \frac{1}{1 + \sigma_\epsilon^2 \left( \sigma_i^{-2} + \frac{1}{\tilde{C}^2} \sigma_x^{-2} \right)} \frac{1}{\sigma_i^2} di. \quad (\text{IA.109})$$

The left-hand-side of this equation is decreasing in  $\tilde{C}$ . The right-hand-side is increasing in  $\tilde{C}$ . If  $\tilde{C} \rightarrow 0$ , the left-hand-side goes to  $\infty$  and the right-hand-side goes to 0. If  $\tilde{C} \rightarrow \infty$ , the left-hand-side goes to 0 and the right-hand-side has a finite positive limit. Therefore, there is a unique solution to the equation, and a unique linear equilibrium.

### C.1.3 Equilibrium elasticities

We now derive demand elasticities. We show how individual demand elasticities respond to the aggregate elasticity. Demand is given by:

$$q_i = \frac{1}{\rho_i} \frac{\hat{\mu}_i - p}{\hat{\sigma}_i^2 + \sigma_\epsilon^2} \quad (\text{IA.110})$$

$$= \frac{1}{\rho_i} \frac{\hat{\sigma}_i^2 \left( \sigma_i^{-2} \mu_i + \frac{B^2}{\tilde{C}^2} \sigma_x^{-2} s(p) \right) - p}{\hat{\sigma}_i^2 + \sigma_\epsilon^2}. \quad (\text{IA.111})$$

Therefore the slope of the demand curve is:

$$\mathcal{E}_i = -\frac{1}{\rho_i} \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \sigma_\epsilon^2} \left( \frac{1}{\tilde{C}^2} \sigma_x^{-2} - \hat{\sigma}_i^{-2} \right) \quad (\text{IA.112})$$

$$= -\frac{1}{\rho_i} \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \sigma_\epsilon^2} \left( \frac{1}{\tilde{C}^2} \sigma_x^{-2} - \sigma_i^{-2} - \frac{1}{\tilde{C}^2} \sigma_x^{-2} \right) \quad (\text{IA.113})$$

$$= \frac{1}{\rho_i} \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \sigma_\epsilon^2} \frac{1}{\sigma_i^2}. \quad (\text{IA.114})$$

Here, we observe clearly the intuition for the role of price informativeness. When prices are more informative, low  $\hat{\sigma}_i^2$ , expected returns respond less to the price, the numerator of the first fraction. However, the perceived risk of the asset also decreases, the denominator of the first fraction. Because of residual uncertainty  $\sigma_\epsilon^2$ , the effect on the asset risk is weaker than the effect on expected returns: the ratio decreases and the trader becomes less aggressive.

More aggregate elasticity leads to more informative prices, so this mechanism will lead to a negative response of individual elasticity to aggregate elasticity. Formally, note that  $\int_i \mathcal{E}_i = \mathcal{E}_{agg} = \tilde{C}^{-1}$ . Plugging in, we obtain:

$$\mathcal{E}_i = \frac{1}{\rho_i} \frac{1}{1 + \sigma_\epsilon^2 \left( \sigma_i^{-2} + \mathcal{E}_{agg}^2 \sigma_x^{-2} \right)} \frac{1}{\sigma_i^2} \quad (\text{IA.115})$$

$$= \frac{1}{\rho_i} \frac{1}{\sigma_i^2 + \sigma_\epsilon^2 + \sigma_i^2 \sigma_\epsilon^2 \sigma_x^{-2} \mathcal{E}_{agg}^2} \quad (\text{IA.116})$$

Clearly, the individual elasticity  $\mathcal{E}_i$  is decreasing in the aggregate elasticity  $\mathcal{E}_{agg}$ . Linearizing this expression, we obtain the counterpart of the degree of strategic response  $\chi > 0$ :

$$\chi = -\frac{d\mathcal{E}_i}{d\mathcal{E}_{agg}} \quad (\text{IA.117})$$

$$= \frac{1}{\rho_i} \frac{2\sigma_i^2\sigma_\epsilon^2\sigma_x^{-2}\mathcal{E}_{agg}}{(\sigma_i^2 + \sigma_\epsilon^2 + \sigma_i^2\sigma_\epsilon^2\sigma_x^{-2}\mathcal{E}_{agg}^2)^2}. \quad (\text{IA.118})$$

## C.2 Price impact

We now consider a model in the style of Kyle (1989), in which investors have non-negligible price impact and take it into account when making trading decisions. This leads to a foundation for a negative degree of strategic response  $\chi$ . Intuitively, when other traders are aggressive, I face a very elastic residual supply curve when sending orders to the market. This implies that my trades will not have a large price impact, hence I can also trade aggressively.

### C.2.1 Price impact

There are  $I$  investors indexed by  $i$ . Each agent has CARA preferences with risk aversion  $\rho_i$ :

$$U_i = \mathbf{E}_i[-e^{-\rho_i W_i}], \quad (\text{IA.119})$$

and initial wealth  $W_i$ . The gross risk-free rate is 1, and the random asset payoff  $f$  is distributed  $\mathcal{N}(\mu, \sigma^2)$ . The asset is in noisy supply  $\bar{x} + x$  with  $\bar{x}$  an exogenous fixed parameter and  $x \sim \mathcal{N}(0, \sigma_x^2)$ .

As in Kyle (1989) we are interested in rational expectation equilibria with imperfect competition. We look for a linear pricing rule  $p = A + Cx$ . We solve for the individual demand strategies and look for linear strategies of the form:

$$d_i = \underline{d}_i - \mathcal{E}_i p \quad (\text{IA.120})$$

### C.2.2 Solving for optimal demand strategies

Investor  $i$  maximizes their profit taking as given the residual demand from other investors' demand schedule. We use market clearing to find the residual supply curve:

$$\begin{aligned} \sum_i d_i &= \bar{x} + x \\ d_i &= \bar{x} + x - \sum_{k \neq i} \underline{d}_k + \left( \sum_{k \neq i} \mathcal{E}_k \right) p \\ p(d_i) &= \underbrace{\left( \sum_{k \neq i} \mathcal{E}_k \right)^{-1}}_{\lambda_{-i}} d_i + \underbrace{\left( \sum_{k \neq i} \mathcal{E}_k \right)^{-1} \cdot \left( \sum_{k \neq i} \underline{d}_k - \bar{x} - x \right)}_{p_{-i}}. \end{aligned} \quad (\text{IA.121})$$

To find the optimal demand of investor  $i$  for the asset, we write their program<sup>8</sup>

$$\begin{aligned} \max_d \mathbf{E}\{f - p(d)|p_{-i}\}d - \frac{\rho_i}{2} \text{Var}\{f - p(d)|p_{-i}\}d^2, \\ \max_d (\mu - p_{-i})d - \lambda_{-i}d^2 - \frac{\rho_i}{2}d^2\sigma^2. \end{aligned} \quad (\text{IA.122})$$

The first order condition and replacing  $p_{-i} = p - \lambda_{-i}d_i$  gives us:

$$d_i = \frac{\mu - p}{\rho_i\sigma^2 + \lambda_{-i}}. \quad (\text{IA.123})$$

We can already see that stronger  $\lambda_{-i}$  leads to less aggressive trading because of a larger price impact. Remember that  $\lambda_{-i}$  is the aggregate of demand elasticities of other investors, a quantity closely related to aggregate elasticity. We now close the equilibrium to show this relation more clearly.

### C.2.3 Solving for aggregate demand elasticity

Given our original demand  $d_i = \underline{d}_i - \mathcal{E}_i p$ , we are able to identify the linear terms as:

$$\underline{d}_i = \frac{\mu}{\rho_i\sigma^2 + \lambda_{-i}}; \quad \mathcal{E}_i = \frac{1}{\rho_i\sigma^2 + \lambda_{-i}} = \frac{1}{\rho_i\sigma^2 + (\mathcal{E}_{agg} - \mathcal{E}_i)^{-1}}, \quad (\text{IA.124})$$

where we define the aggregate elasticity:

$$\mathcal{E}_{agg} = \sum_i \mathcal{E}_i. \quad (\text{IA.125})$$

Next we show that there is a unique solution for the aggregate elasticity. From the expression in equation (IA.124), we remark that  $\mathcal{E}_i$  solves a quadratic equation. We rule out the larger of the two roots and the solution is<sup>9</sup>

$$\mathcal{E}_i = \frac{1}{2} \left( \frac{2}{\rho_i\sigma^2} + \mathcal{E}_{agg} - \sqrt{\left(\frac{2}{\rho_i\sigma^2}\right)^2 + \mathcal{E}_{agg}^2} \right) \quad (\text{IA.126})$$

To show that there is a unique stable equilibrium we consider the fixed point problem  $F(x) = x$ , with  $F$  defined by:

$$f_i(x) = \frac{1}{2} \left( \frac{2}{\rho_i\sigma^2} + x - \sqrt{\left(\frac{2}{\rho_i\sigma^2}\right)^2 + x^2} \right), \quad (\text{IA.127})$$

$$F(x) = \sum_i f_i(x). \quad (\text{IA.128})$$

<sup>8</sup>Note that expectation and variances are conditional on the residual demand curve  $p_{-i}$ , which is equivalent to conditioning on  $p$ .

<sup>9</sup>The larger root is such that  $\mathcal{E}_i > \mathcal{E}_{agg}$  which violates  $\sum_i \mathcal{E}_i = \mathcal{E}_{agg}$ .

The function  $F$  is positive, increasing, and concave. Moreover  $F(0) = 0$ ,  $F'(0) = I/2$ , and  $\lim_{x \rightarrow +\infty} F'(x) = 0$ , we conclude that there is a unique non-zero solution for  $\mathcal{E}_{agg}$  as long as  $I \geq 3$ .

The relation derived in (IA.126) between  $\mathcal{E}_i$  and  $\mathcal{E}_{agg}$  is not linear. We can approximate this equation linearly by  $\mathcal{E}_i = \underline{\mathcal{E}}_i - \chi \mathcal{E}_{agg}$  with

$$\chi = -\frac{1}{2} \left( 1 - \frac{\mathcal{E}_{agg}}{\sqrt{\mathcal{E}_{agg}^2 + \left(\frac{2}{\gamma_i \sigma^2}\right)^2}} \right) < 0 \quad (\text{IA.129})$$

This expression gives bounds on the value of  $\chi$ :  $-1/2 \leq \chi < 0$ .

### C.3 Imperfect information

Assume your optimal elasticity is  $\mathcal{E}_i = \underline{\mathcal{E}}_i - \chi \mathbf{E}[\mathcal{E}_{agg} | \mathcal{F}_i]$ . With perfect knowledge, you obtain:  $\mathcal{E}_i = \underline{\mathcal{E}}_i - \chi \mathcal{E}_{agg}$ . Assume the agent observes a signal  $\hat{\mathcal{E}}_{agg} = \mathcal{E}_{agg} + \epsilon$  (with variance  $\sigma_\epsilon^2$ ) and has a prior  $\mathcal{E}_{agg} \sim \mathcal{N}(\bar{\mathcal{E}}, \sigma^2)$ . Then we have:

$$\mathbf{E}[\mathcal{E}_{agg} | \hat{\mathcal{E}}_{agg}] = \frac{1}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma^2}} \left( \frac{1}{\sigma_\epsilon^2} \hat{\mathcal{E}}_{agg} + \frac{1}{\sigma^2} \bar{\mathcal{E}} \right). \quad (\text{IA.130})$$

Therefore:

$$\mathcal{E}_i = \underline{\mathcal{E}}_i - \chi \frac{1}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma^2}} \frac{1}{\sigma^2} \bar{\mathcal{E}} - \chi \underbrace{\frac{\frac{1}{\sigma_\epsilon^2}}{\frac{1}{\sigma_\epsilon^2} + \frac{1}{\sigma^2}}}_{<1} \hat{\mathcal{E}}_{agg}. \quad (\text{IA.131})$$

### C.4 Partial equilibrium thinking

We repeat the calculation of Section 2.3 on the rise in passive investing in a situation with partial equilibrium thinking.

We assume that all investors are homogenous and their initial elasticity is  $\mathcal{E}_i = \mathcal{E}_0$ . What happens to the economy if a fraction  $1 - \alpha$  of these investors become passive? Their elasticity reduces to zero. To model partial equilibrium thinking, we assume that active investors only react to the effect of the switch to passive on aggregate elasticity and do not take into account the collective response of other active investors.

Because investors are infinitesimal, this corresponds to forecasting a change in elasticity of  $\Delta \mathcal{E}_{agg}^{\text{forecast}} = -(1 - \alpha) \mathcal{E}_0$ . This implies that each active investor changes her elasticity by

$$\Delta \mathcal{E}_i = -\chi \Delta \mathcal{E}_{agg}^{\text{forecast}} = \chi(1 - \alpha) \mathcal{E}_0. \quad (\text{IA.132})$$

Aggregating across all investors, the new aggregate elasticity is:

$$\mathcal{E}_{NEW}^{PET} = \alpha \mathcal{E}_0 (1 + \chi(1 - \alpha)) = \alpha \mathcal{E}_0 + (1 - \alpha) \chi \alpha \mathcal{E}_0. \quad (\text{IA.133})$$

The new elasticity with partial equilibrium thinking is in contrast to our baseline model in Section 2.3,  $\mathcal{E}_{NEW}$ :

$$\mathcal{E}_{NEW}^{PET} - \mathcal{E}_{NEW} = (1 - \alpha)\mathcal{E}_0\alpha\chi\frac{\alpha\chi}{1 + \alpha\chi}. \quad (\text{IA.134})$$

The difference is positive when  $\chi > 0$ . Because investors do not account for the response of others, they overreact to the initial change in elasticity. This leads to a relatively higher final level of aggregate elasticity and is therefore akin to a larger degree of strategic response. For some parameter values, the response is so strong that it flips the sign of the change in aggregate elasticity.

## D Data

### D.1 Institutional holdings data

Institutional investment managers with \$100 million or more in assets under their investment discretion are required to disclose their ownership of Section 13(f) securities as of the end of each calendar quarter to the SEC within 45 days after the end of the calendar quarter. The filing requirement applies to both U.S. domestic investment managers, and, under certain conditions regarding the course of their business, foreign investment managers. The official list of 13(f) securities is made available by the SEC shortly after each quarter end. It primarily includes U.S. exchange-traded stocks, shares of closed-end investment companies, and shares of exchange-traded funds.

We obtain data on 13F filings from 2001Q1 until 2017Q4 from Backus, Conlon, and Sinkinson (2020). Mirroring their approach, we extend their sample until 2020Q4. To do so we start with a SEC linking table, which provides a list with links to all 13F-HR and 13F-HR/A filings of a given quarter. Based on those links, we scrape all filings, and subsequently parse the filings based on a Perl script generously provided by Backus, Conlon, and Sinkinson (2019). The script corrects common filing issues in 13F filings. Finally, we merge the scraped 13F data with CRSP and Compustat.

### D.2 CRSP and Compustat

We obtain market capitalization data for stocks from CRSP and apply standard filters: we keep stocks traded on NYSE, NASDAQ and AMEX, and filter to ordinary common shares with share codes 10, 11, 12 and 18 in CRSP. These stocks make up the universe of all assets in our model.

Additional stock-level information comes from quarterly and annual Compustat files. In particular, we closely follow Kojien and Yogo (2019) and their data definitions for building a set of stock characteristics: book equity, investment (defined as growth in total assets), operating profitability (as defined in Fama and French (2015)), and dividend yield as a fraction of book equity. Characteristics are winsorized at the 2.5% and the 97.5% level each quarter. Characteristics that are denominated in dollar values, such as market equity and book equity, are denominated in million dollars.

We match CRSP and Compustat based on the standard linking table on WRDS. Finally, we use CUSIP identifier information from CRSP to merge the CRSP-Compustat merged stock-level data to 13F holdings. We exclude stocks for which institutional ownership is greater than 100% based on the 13F data.

### D.3 Measuring passive investing

Passive investors are insensitive to prices. At each date, we identify passive investors as investors with elasticity close to zero in a Kojien and Yogo (2019) type demand system:

$$\log \frac{w_{ik}}{w_{i0}} - p_k = \underline{d}_{0i} + \underline{d}'_{1it} X_k^{(d)} - \mathcal{E}_i^{fixed} p_k + \epsilon_{ik}, \quad (\text{IA.135})$$



where  $X^{(d)}$  contains log book equity. An investor is defined as passive if their fixed elasticity is close to zero, i.e.  $\mathcal{E}_i^{fixed} < \kappa$ , for small  $\kappa$ . We choose  $\kappa = 0.06$  to calibrate the level of passive investing, and define a stock's passive share as the ownership-weighted average of an indicator that is 1 if the investor is passive, and zero otherwise. That is,

$$|Active_k| \equiv 1 - \sum_i \underbrace{\frac{w_{ik}A_i}{\exp(p_k)}}_{\text{ownership share}} 1_{\{\mathcal{E}_i^{fixed} < \kappa\}}. \quad (\text{IA.136})$$

Appendix Figure IA.8 shows the the cross-sectional median of  $|Active_k|$  over time. We validate our measure in the setting of Russell index switching. Specification (4) of Appendix Table IA.4 shows that our measure of passive investing in the cross-section strongly responds to stocks switching between the Russell 1000 and 2000 indices, by about 4% of total ownership.

## D.4 Additional data definitions

There are a number of additional data steps that define the final estimation sample.

**Defining the outside asset.** For the logit demand system we define any stock with missing stock characteristics or CRSP share code 12 or 18 in a given quarter as part of the outside asset for that particular quarter. Of the remaining stocks, any stock with fewer than 20 investment managers invested in it is also part of the outside asset.

**Defining the household sector.** Investment managers with fewer than 100 stocks in their portfolio are filtered out, such that their assets are part of the residual household sector. The residual household sector contains direct household holdings, but also an amalgamation of holdings from small investment managers with AUM below the reporting threshold, certain foreign investors, and investment managers with fewer than 100 stocks in their portfolio. As in Kojien and Yogo (2019), this residual household sector is modeled as one investor in the demand system, to ensure that the number of shares held adds up to the number of shares outstanding.

**Measuring the investment universe.** We define any stock that an investment manager has held over the past three years as part of her investment universe. This follows Kojien and Yogo (2019), who show that the measured investment universe using this approach is very stable over time. The investment universe is primarily used during the estimation procedure to construct our instruments for a stock's market equity as in equation (21), and a stock's aggregate demand elasticity as in equation (22).

**Pooling investors during estimation.** For our baseline estimation, we pool together investors that hold fewer than 1,000 stocks in a quarter and are classified as active. Investors are grouped based on their assets under management, with the number of groups chosen such that on average, each group holds 2,000 stocks. Specifically, we assume that all

investors within the same group have the same demand parameters, except for the constant of equation (IA.151), which we leave as investor-specific (for example, this absorbs variation in quantity of outside asset  $w_{i0}$ ). That is, we estimate equation (IA.151) at the group level with an investor fixed effect.

**Weighting investors during estimation.** For the pooled regression in equation (IA.152), we weight each observation such that each investor-group contributes equally to the regression. That is, each observation receives a weight  $1/|\mathcal{K}_i|$ . For example, consider a simplified example with two investors and two assets. Investor  $A$  holds both assets, while investor  $B$  holds only one of the assets. We would assign weight of 0.5 to each position of investor  $A$ , and a weight of 1 to the observation for investor  $B$ . The estimate of  $\chi$  is robust to different weighting schemes, as shown in Table 2.

# E Identification Strategy

## E.1 Moment conditions

We estimate the model using the method of moments. All of the moment conditions derive from the identifying assumption of equation (23). We list these moments here:

$$\mathbf{E} [\epsilon_{jk} \mathbf{1}_{\{j=i\}}] = 0, \forall i \quad (\text{IA.137})$$

$$\mathbf{E} [\epsilon_{jk} X_k^{(d)} \mathbf{1}_{\{j=i\}}] = 0, \forall i \quad (\text{IA.138})$$

$$\mathbf{E} [\epsilon_{jk} \hat{p}_{k,i} \mathbf{1}_{\{j=i\}}] = 0, \forall i \quad (\text{IA.139})$$

$$\mathbf{E} [\epsilon_{jk} X_k^{(e)} \hat{p}_{k,j} \mathbf{1}_{\{j=i\}}] = 0, \forall i \quad (\text{IA.140})$$

$$\mathbf{E} [\epsilon_{jk} \hat{\mathcal{E}}_{agg,k}] = 0 \quad (\text{IA.141})$$

$$\mathbf{E} [\epsilon_{jk} \hat{\mathcal{E}}_{agg,k} \hat{p}_{k,j}] = 0 \quad (\text{IA.142})$$

There are exactly as many moment conditions as model parameters.

## E.2 Solving the reflection problem

One challenge for identification is the reflection problem. How can we separate the individual component of demand elasticity from the strategic response to other investors? We show that the presence of variation in investor population across stocks allows to solve this problem. To isolate this argument from other identification concerns, we assume that we observe individual elasticities,  $\mathcal{E}_{ik}$ . For exposition purposes, we focus on a simplified version of the model in which  $\underline{\mathcal{E}}_i$  does not depend on asset characteristics.

We provide sufficient conditions for the uniqueness of a decomposition of the individual elasticities into investor-specific elasticities  $\underline{\mathcal{E}}_i$  and the strategic response controlled by  $\chi$ . After proving this result, we come back to the economic content and the empirical relevance of these conditions.

Before stating the theorem, we introduce a few notations. We define the undirected graph  $\mathcal{G}$  of investor-stock connections. The vertices and the nodes are the investors  $i$  and the stocks  $k$ . There is an edge between  $i$  and  $k$  if and only if  $i \in I_k$ . There are no edges between two investors or two stocks.

**Theorem 2.** *A decomposition of demand elasticities  $\{\mathcal{E}_{ik}\}_{i,k}$  into individual elasticities  $\{\underline{\mathcal{E}}_i\}_i$  and the degree of strategic response  $\chi$  is unique if:*

- (a) *The graph  $\mathcal{G}$  of investor-stock connections is connected.*
- (b) *Position-weighted averages of demand elasticities are not constant across stocks: there exists  $k$  and  $k'$  such that  $\sum_{i \in I_k} w_{ik}/p_k A_i \underline{\mathcal{E}}_i \neq \sum_{i \in I_{k'}} w_{ik'}/p_{k'} A_i \underline{\mathcal{E}}_i$ .*

*Proof.* Let us assume that there exist two distinct decompositions  $(\{\underline{\mathcal{E}}_i^{(1)}\}_i, \chi^{(1)}) \neq (\{\underline{\mathcal{E}}_i^{(2)}\}_i, \chi^{(2)})$  and the two conditions (a) and (b) hold. Each decomposition for  $l \in \{1, 2\}$

satisfies the two conditions of the elasticity layer

$$\mathcal{E}_{agg,k} = \sum_{i \in I_k} w_{ik} A_i / p_k \mathcal{E}_{ik}, \quad \text{for all } k \in \mathcal{K} \quad (\text{IA.143})$$

$$\mathcal{E}_{ik} = \underline{\mathcal{E}}_i^{(l)} - \chi^{(l)} \mathcal{E}_{agg,k}, \quad \text{for all } k \in \mathcal{K} \text{ and } i \in I_k. \quad (\text{IA.144})$$

We subtract the decomposition of  $\mathcal{E}_{ik}$  for  $l = 1$  from the decomposition for  $l = 2$  and obtain:

$$(\chi^{(2)} - \chi^{(1)}) \mathcal{E}_{agg,k} = \underline{\mathcal{E}}_i^{(2)} - \underline{\mathcal{E}}_i^{(1)}, \quad \text{for all } k \in \mathcal{K} \text{ and } i \in I_k. \quad (\text{IA.145})$$

Here we see immediately that if  $\chi^{(1)} = \chi^{(2)}$ , then for all  $i$ ,  $\underline{\mathcal{E}}_i^{(1)} = \underline{\mathcal{E}}_i^{(2)}$ , thus violating the initial assumption of distinct decompositions. Hence, we focus on the case of  $\chi^{(1)} \neq \chi^{(2)}$ .

We define the function:

$$f(x) = \begin{cases} (\chi^{(2)} - \chi^{(1)}) \mathcal{E}_{agg,x} & \text{for } x \in \mathcal{K} \\ \underline{\mathcal{E}}_x^{(2)} - \underline{\mathcal{E}}_x^{(1)} & \text{for } x \in I. \end{cases} \quad (\text{IA.146})$$

We restate the equality of equation (IA.145) as:

$$f(x) = f(x'), \quad \text{if and only if there is an edge between } x \text{ and } x' \text{ on } \mathcal{G}. \quad (\text{IA.147})$$

Therefore, since the graph  $\mathcal{G}$  is connected:  $\forall x, x', f(x) = f(x')$ , and  $f$  is a constant. We write the constant  $f = a$ , and plug in the constant in the aggregation of individual elasticities:

$$\mathcal{E}_{agg,k} = \sum_{i \in I_k} w_{ik} A_i / p_k \mathcal{E}_{ik} = \sum_{i \in I_k} w_{ik} A_i / p_k \underline{\mathcal{E}}_i^{(1)} - \chi^{(1)} \sum_{i \in I_k} w_{ik} A_i / p_k \mathcal{E}_{agg,k} \quad (\text{IA.148})$$

$$\iff (1 + \chi^{(1)}) \mathcal{E}_{agg,k} = \sum_{i \in I_k} w_{ik} A_i / p_k \underline{\mathcal{E}}_i^{(1)} \quad (\text{IA.149})$$

$$\iff (1 + \chi^{(1)}) \frac{a}{\chi^{(2)} - \chi^{(1)}} = \sum_{i \in I_k} w_{ik} A_i / p_k \underline{\mathcal{E}}_i^{(1)} \quad \text{for all } k, \quad (\text{IA.150})$$

where we use  $\mathcal{E}_{agg,k} = a / (\chi^{(2)} - \chi^{(1)})$ . Equation (IA.150) violates assumption (b), which concludes the proof.  $\blacksquare$

The intuition behind theorem 2 is that identification relies on comparing the behavior of one investor for two different stocks with different populations of investors. If this investor trades less aggressively when surrounded by more aggressive investors, we conclude that the degree of strategic response  $\chi$  is positive. A challenge to implement this comparison is that we already need to know the elasticity of these other investors. This is a chicken-and-egg question. The ability to find a unique solution to this problem relies on being able to cycle through investors with enough variation in composition: this is the essence of conditions (a) and (b).

To better understand why these conditions are important, we show examples of how the model is not identified when either (a) or (b) is violated. Starting with (a), let us consider the case where each stock has its own non-overlapping population of investors. In this case, there is no identification. Because a given investor only invests in one stock, it

is not possible to tell if this investor is aggressive because of her own characteristics or in response to the other investors. As an example that violates condition (b), consider the case in which all investors have the same size and relative portfolio positions such that:  $\forall k, k', w_{ik}A_i/p_k = w_{ik'}A_i/p_{k'}$ . Investor composition is the same for all stocks and therefore there is no information in comparing different stocks. Relatedly, we could also consider a violation of (b) where all individual elasticities are identical across investors:  $\underline{\mathcal{E}}_i = \underline{\mathcal{E}}$ . Then, for all  $k$  we have  $\sum_{i \in I_k} w_{ik}/p_k A_i \underline{\mathcal{E}}_i = \underline{\mathcal{E}}$ : the aggregate elasticity for all stocks is identical. Intuitively, even though there is variation in investor composition across stocks, all investors behave the same way in terms of elasticity. This is equivalent to having a single investor, and we cannot separate individual elasticities from the response to other investors.

How can we assess these conditions empirically? The graph  $\mathcal{G}$  of investors-stocks connections can be observed directly in our data and we can assess immediately that condition (a) is satisfied using known algorithms such as depth-first-search. Condition (b) is potentially more challenging because it relies on parameter estimates  $\underline{\mathcal{E}}_i$ . However, inspecting the condition shows it holds generically. Condition (b) stipulates the equality of  $K$  linear forms applied to the vector  $(\underline{\mathcal{E}}_i)_i$ . It is violated if and only if  $(\underline{\mathcal{E}}_i)_i \in \bigcap_{k>1} (w_k - w_1)^\perp$ , a set of measure 0 for almost all combinations of  $w_k$ . In addition, there is still the possibility of verifying whether the condition is satisfied empirically, once the econometrician has found a set of parameter estimates.

### E.3 Numerical procedure

We describe our estimation procedure, which solves a series of nested problems.

**Step 1.** Given a guess for  $(\chi, \xi)$  and  $\{\mathcal{E}_{agg,k}\}_k$ , we can estimate all remaining model parameters by two-stage least squares regression investor by investor. This corresponds to estimating the following regression for each investor  $i$ :

$$\log \frac{w_{ik}}{w_{i0}} - p_k - \chi \mathcal{E}_{agg,k} p_k - \xi \mathcal{E}_{agg,k} = \underline{d}_{0i} + \underline{d}'_{1it} X_k^{(d)} - \left( \underline{\mathcal{E}}_{0i} + \underline{\mathcal{E}}'_{1i} X_k^{(e)} \right) p_k + \epsilon_{ik}, \quad (\text{IA.151})$$

where  $p_k$  and  $X_k^{(e)} p_k$  are instrumented by  $\hat{p}_{k,i}$  and  $X_k^{(e)} \hat{p}_{k,i}$ . Estimating these regressions is equivalent to solving the moment conditions (IA.137) to (IA.140).

**Step 2.** Given a guess for  $(\chi, \xi)$ , we look for equilibrium values of  $\{\mathcal{E}_{agg,k}\}_k$ . We start from the aggregate elasticities implied by the model of Kojien and Yogo (2019). We run step 1 above. With the newly estimated  $\underline{\mathcal{E}}_i$  and the parameter  $\chi$ , we solve explicitly for the equilibrium elasticity they imply by solving the linear system of equations (14) and (18). We update our guessed aggregate elasticity by taking a weighted average of the previous iteration and these new implied values with weights of 75% and 25%. We repeat this updating process until the values of  $\{\mathcal{E}_{agg,k}\}_k$  converge. This step ensures that our estimated model satisfies the 2-layer equilibrium.

**Step 3.** We estimate  $\chi$  and  $\xi$ . We start from a guess for  $(\chi, \xi)$  and run step 2 to find the aggregate elasticities it implies. With these values, we estimate the pooled regression of

equation (19):

$$\log \frac{w_{ik}}{w_{i0}} - p_k = \underline{d}_{0i} + \underline{d}'_{1it} X_k^{(d)} + \xi \mathcal{E}_{agg,k} - \left( \underline{\mathcal{E}}_{0i} + \underline{\mathcal{E}}'_{1i} X_k^{(e)} - \chi \mathcal{E}_{agg,k} \right) p_k + \epsilon_{ik}, \quad (\text{IA.152})$$

using two-stage least squares with all the instruments of the investor-level regression,  $\hat{\mathcal{E}}_{agg,k}$  and  $\hat{\mathcal{E}}_{agg,k} \hat{P}_{k,i}$ . This is a very large scale regression with many fixed effects and investor-specific coefficients. We speed up the estimation of this large-scale regression tremendously by taking advantage of the Frisch-Waugh-Lovell theorem. We absorb all individual-level variables using investor-specific regressions, and are left with only the coefficients  $\chi$  and  $\xi$  to estimate in the pooled data.

Define as  $f(\cdot)$  the function that maps the guess  $(\chi, \xi)$  to estimates for  $\chi$  and  $\xi$  in the pooled regression, and define the fixed point function  $F(\chi, \xi) \equiv f(\chi, \xi) - (\chi, \xi)$  as the difference between the estimates from the pooled regression and the guess.

**Step 4.** The pooled regression gives  $F(\chi, \xi)$ . We use a bivariate quasi-Newton method, in which we approximate the Jacobian of  $F$  numerically via finite differences by varying  $\chi$  and  $\xi$  by a small  $\epsilon$ , to find a root of  $F(\chi, \xi)$ , which constitutes a fixed point for  $(\chi, \xi)$ . With such a fixed point, we are sure that our estimates satisfy simultaneously all the moment conditions of Appendix Section E.1 and the 2-layer equilibrium.

Algorithm E.1 summarizes the numerical procedure to obtain a fixed point for  $(\chi, \xi)$  in pseudo-code.

---

**Algorithm E.1: Numerical procedure solving for a fixed point of  $(\chi, \xi)$ .**

---

```

1  begin
2  Initialize starting values  $(\chi^{(0)}, \xi^{(0)})$ 
3   $h \leftarrow 0$ 
4  while  $(\|F(\chi^{(h-1)}, \xi^{(h-1)})\| > \text{tol})$  or  $(h = 0)$ 
5  Initialize  $\{\mathcal{E}_{agg,k}^{(0)}\}_k$  at  $\{\mathcal{E}_{fixed,k}\}_k$ 
6  for  $n$  in  $1:N$ 
7  Update investor-specific parameters conditional on  $\{\mathcal{E}_{agg,k}^{(n-1)}\}_k$  and  $(\chi^{(h)}, \xi^{(h)})$  (Step 1).
8  Aggregate to determine  $\{\mathcal{E}_{agg,k}^{(n)}\}_k$  conditional on  $(\chi^{(h)}, \xi^{(h)})$  (Step 2).
9  end
10 Determine  $f(\chi^{(h)}, \xi^{(h)})$ , i.e. estimate  $(\chi, \xi)$  conditional on  $\{\mathcal{E}_{agg,k}^{(N)}\}_k$  (Step 3).
11  $F(\chi^{(h)}, \xi^{(h)}) \leftarrow f(\chi^{(h)}, \xi^{(h)}) - (\chi^{(h)}, \xi^{(h)})$ 
12  $\hat{J}(\chi^{(h)}, \xi^{(h)}) \leftarrow \frac{1}{\epsilon} (F(\chi^{(h)} + \epsilon, \xi^{(h)}) - F(\chi^{(h)}, \xi^{(h)}), F(\chi^{(h)}, \xi^{(h)} + \epsilon) - F(\chi^{(h)}, \xi^{(h)}))$ 
13  $(\chi^{(h+1)}, \xi^{(h+1)}) \leftarrow (\chi^{(h)}, \xi^{(h)}) - \hat{J}^{-1}(\chi^{(h)}, \xi^{(h)}) F(\chi^{(h)}, \xi^{(h)})$  (Step 4)
14  $h \leftarrow h + 1$ 
15 end
16 return  $(\chi^{(h)}, \xi^{(h)})$ 
17 end

```

---

Lines 2 and 3 initialize the numerical procedure. Starting values  $(\chi^{(0)}, \xi^{(0)})$  are based on past experience with the algorithm as the Newton method may fail for starting values too far removed from a root of  $F$ . Line 4 starts a while loop that ends when a solution is found, i.e. when the norm of  $F$  is below some small tolerance level. Lines 5 to 9 solve for an elasticity equilibrium conditional on the current iteration  $(\chi^{(h)}, \xi^{(h)})$ : First, we initialize aggregate elasticities based on a fixed elasticity model (IA.135). Then, we iterate back and

forth between estimating investor-specific parameters conditional on aggregate elasticities and aggregating individual elasticities until an equilibrium is found. In line 10 we estimate the pooled regression of equation (IA.152) conditional on  $\{\mathcal{E}_{agg,k}^{(N)}\}_k$ . Line 11 updates the fixed point function. Line 12 approximates the Jacobian of  $F$  at  $(\chi^{(h)}, \xi^{(h)})$  via a finite difference approach. Line 13 updates  $\chi$  and  $\xi$  via a Newton step, and line 14 increases the iterator.

## F Trading Big and Small Stocks

We investigate whether firms trade big and small stocks differently. Our estimates of elasticities by stocks suggest that the demand for large stocks is more inelastic (see Figure 3). To explain this result, one hypothesis is that large stocks mechanically tend to receive a high portfolio weight and that investors are unwilling to adjust their largest positions. For example, a 10% relative increase in portfolio weight would create much larger tracking error to the index for large positions than for small positions. Also, the granular nature of large stocks imply that they have fewer substitutes.

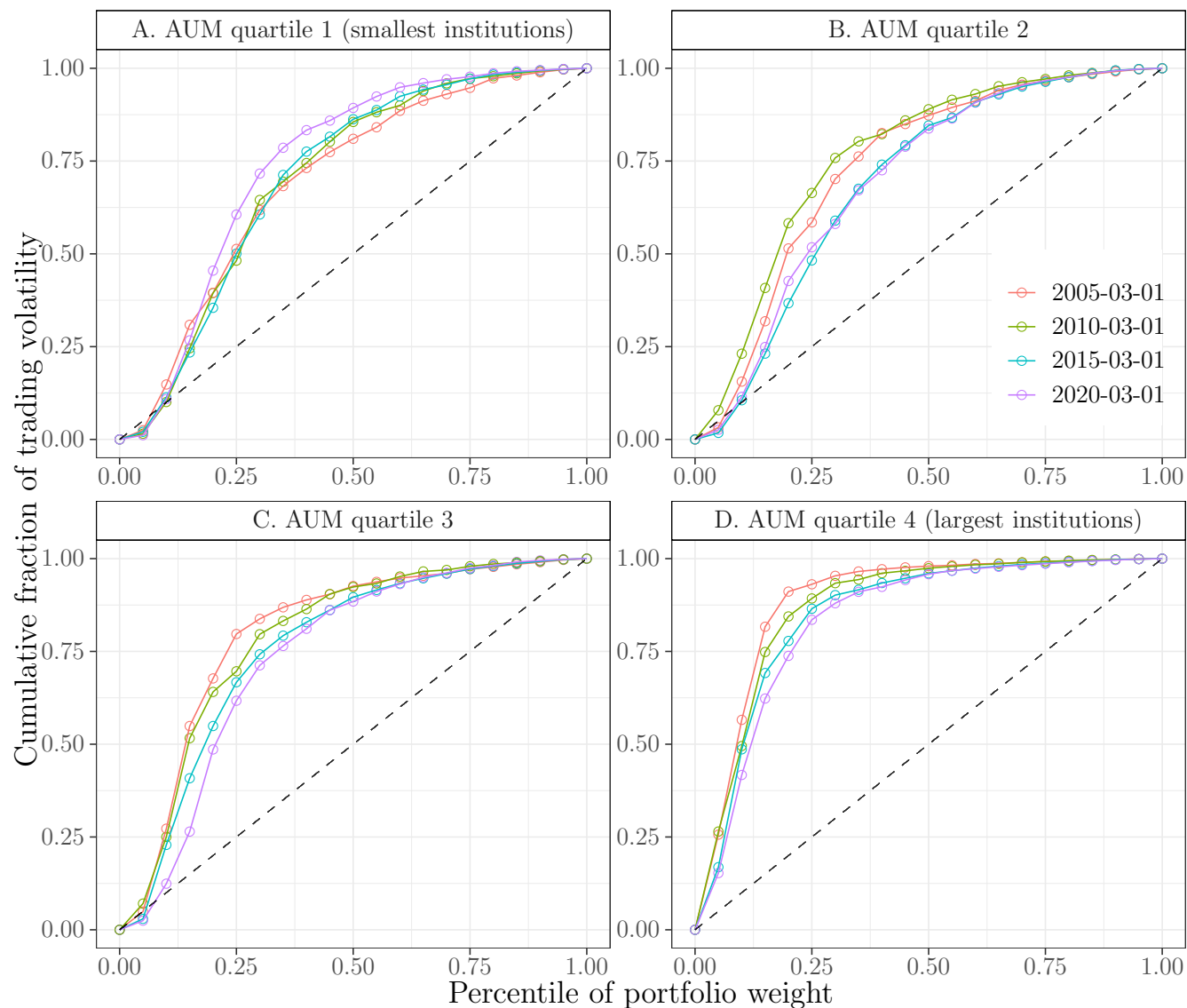
To complement our structural results and investigate this hypothesis, we compare the trading activity of investors across the distribution of their portfolio. For a given investor-quarter, we compute for each stock the squared relative change in the number of shares:

$$\text{Trading Activity}_{i,k,t} = \left[ \left( \frac{A_{i,t}w_{ik,t}}{p_{k,t}} - \frac{A_{i,t-1}w_{ik,t-1}}{p_{k,t-1}} \right) / \frac{A_{i,t}w_{ik,t}}{p_{k,t}} \right]^2 \quad (\text{IA.153})$$

We sort positions by portfolio weights, and compute the ratio of the cumulative sum of trading activity to the total sum. This gives us a relation between the percentile of portfolio weight and the cumulative share of total trading activity. We average this relation within size groups of investors and present our results in Figure IA.1 for various dates.

If trading activity is as intense for all portfolio weights, this curve should coincide with the 45-degree line. Instead, we see that the curve is always above the 45-degree line and particularly flat along the largest investor positions. This implies that there is relatively less trading activity for the largest stocks. In addition, we observe that this pattern is more pronounced for the largest investors (panel D) than for small investors (panel A). Because larger investors are more important for the biggest stocks, this will amplify the lack of trading activity for the biggest stocks.





**Figure IA.1. Trading activity across portfolio positions.** Figure IA.1 presents the cumulative share of trading activity (defined in equation (IA.153)) by quantiles of investor portfolio weights. We aggregate the statistics by date and quartiles of assets under management.

## G Appendix Tables

**Table IA.1. Estimates of the degree of strategic response  $\chi$  under alternative specifications starting 2003Q3**

	Estimates for $\chi$		
	Median	25th pct.	75th pct.
(1) Baseline Specification	2.21	1.85	2.75
(2) BE-weighted Instrument for $\mathcal{E}_{agg}$	1.91	1.54	2.29
(3) Additional Controls	2.49	2.09	3.28
(4) AUM-weighted Regression	2.3	1.86	2.77
(5) Book-weighted Regression	2.24	1.77	2.68
(6) Investor-Type Grouping	2.46	1.96	2.93
(7) Constant $\chi$	2.07		
(8) No Instrument for $\mathcal{E}_{agg}$	1.27	0.88	1.59
(9) No Instruments	1.01	0.73	1.41

Table IA.1 presents statistics of estimates of  $\chi$  across dates (2003Q3–2020Q4) under various specifications. Our baseline specification (1) estimates  $\chi$  given aggregate elasticities  $\mathcal{E}_{agg,k}$  each period via the regression:

$$\log \frac{w_{ik}}{w_{i0}} - p_k = \underline{d}_{0i} + \underline{d}'_{1i} X_k^{(d)} + \xi \mathcal{E}_{agg,k} - \left( \underline{\mathcal{E}}_{0i} + \underline{\mathcal{E}}'_{1i} X_k^{(e)} - \chi \mathcal{E}_{agg,k} \right) p_k + \epsilon_{ik},$$

where  $X_k^{(d)}$  contains log book equity and log book equity squared.  $X_k^{(e)}$  is log book equity. Active investors with fewer than 1,000 stock holdings are grouped together based on their assets under management such that each group on average contains 2,000 stock holdings. The regression is weighted such that each group's weights sum to the same constant. Specification (2) shows estimates of  $\chi$  based on the book-equity weighted instrument. Specification (3) adds additional characteristics to  $X_k^{(d)}$ ; profitability, investment and dividends relative to book equity. Specification (4) value-weights the regression by weighting investors by their AUM. Specification (5) similarly value-weights the regression by weighting investors by their book assets. Specification (6) groups investors both by investor type and AUM. Institutional investors whose type we cannot determine are bundled together in a separate group. Specification (7) imposes for  $\chi$  to be constant across time in the estimation, with each year receiving equal weight. Specification (8) reports results without instrumenting for the aggregate elasticity  $\mathcal{E}_{agg}$ . Specification (9) additionally removes the instrument for prices.

**Table IA.2. Summary statistics of aggregate elasticity  $\mathcal{E}_{agg}$  with book-equity weighted instrument**

Panel A: Statistics of average elasticity across stocks				
	Average	25th pct.	Median	75th pct.
Elasticity $\mathcal{E}_{agg}$	0.504	0.428	0.482	0.569
Fixed elasticity	0.389	0.357	0.389	0.442
Panel B: Regression coefficient (by dates) of elasticity on size				
	Average	25th pct.	Median	75th pct.
Elasticity $\mathcal{E}_{agg}$	-0.111	-0.12	-0.104	-0.091
Fixed elasticity	-0.0286	-0.0309	-0.0272	-0.0249
Panel C: Residual cross-sectional standard deviation of elasticity				
	Average	25th pct.	Median	75th pct.
Elasticity $\mathcal{E}_{agg}$	0.0548	0.0438	0.05	0.0591
Fixed elasticity	0.0842	0.0739	0.0826	0.0917

Table IA.2 presents statistics of the aggregate elasticity  $\mathcal{E}_{agg,k,t}$  using the book-equity weighted instrument. We estimate the elasticities in our baseline model and in a specification with fixed elasticities ( $\chi = 0$  as in Kojien and Yogo (2019)). Panel A has summary statistics of the average elasticity by date. Panel B shows summary statistics of the coefficient  $\beta_t$  from the regression  $\mathcal{E}_{agg,k,t} = \alpha_t + \beta_t p_{k,t} + \varepsilon_{k,t}$  by date. Panel C reports summary statistics of the cross-sectional standard deviation of the residual from the regression described in Panel B. The sample period is 2001–2020.

**Table IA.3. Change in aggregate stock-level elasticity  $\mathcal{E}_{agg,k}$  on the active share using estimates from the model with a constant  $\chi$**

	Change in Elasticity				
	(1)	(2)	(3)	(4)	(5)
Change in Active share	0.321*** (0.058)	0.462*** (0.028)	0.441*** (0.026)	0.412*** (0.025)	0.428*** (0.049)
Date Fixed Effects		Yes	Yes	Yes	Yes
Stock Fixed Effects			Yes		
Controls				Yes	Yes
Estimator	OLS	OLS	OLS	OLS	IV
$N$	50,292	50,292	49,661	50,292	10,619
$R^2$	0.017	0.269	0.282	0.318	0.627
First-stage $F$ statistic					9.444
First-stage $p$ value					0.000

Table IA.3 reports a panel regression of annual log change in stock level elasticity  $\mathcal{E}_{agg,k}$  on the annual log change in the active share  $|Active_k|$ . We use the estimates from the model with a constant value of  $\chi$  over time. Column 2 adds date fixed effects. Column 3 adds stock fixed effects. Column 4 uses date fixed effects and controls for lagged book equity and annual log changes of log book equity. Column 5 instruments the log change in the active share  $|Active_k|$  between Q1 and Q2 in any given year by two indicator variables corresponding to stocks switching between Russell 1000 and 2000 in either direction. In this column, the sample is restricted to stocks with CRSP market capitalization ranked between 500 to 1500 as of the end of Q1. The sample period is 2001–2020 for columns 1-4, and 2007–2020 for column 5. Standard errors are 2-way clustered by date and stock for columns 1-4, and clustered by date for column 5.

**Table IA.4. Change in aggregate stock-level elasticity  $\mathcal{E}_{agg,k}$  on the active share around Russell index reconstitution**

	Log Change in Elasticity			Log Change in Active share
	(1)	(2)	(3)	(4)
Log Change in Active share	0.389*** (0.066)	0.431*** (0.021)		
Log Change in Book Equity	-1.158*** (0.049)	-1.159*** (0.049)	-1.151*** (0.050)	0.018 (0.018)
Lagged Book Equity	0.005 (0.007)	0.005 (0.007)	0.007 (0.007)	0.003 (0.002)
Switch from Russell 2000 to 1000			0.017*** (0.004)	0.035*** (0.008)
Switch from Russell 1000 to 2000			-0.011 (0.006)	-0.043*** (0.012)
Date Fixed Effects	Yes	Yes	Yes	Yes
Estimator	IV	OLS	OLS	OLS
$N$	10,619	10,619	10,619	10,619
$R^2$	0.748	0.748	0.670	0.090
First-stage $F$ statistic	9.444			
First-stage $p$ value	0.000			

Table IA.4 reports a panel regression of log change in stock level elasticity  $\mathcal{E}_{agg,k}$  on the log change in the active share  $|Active_k|$  around Russell index reconstitution, in particular between end of Q2 and Q1, from 2007 to 2020. The sample is restricted to stocks with CRSP market capitalization ranked between 500 to 1500 as of the end of Q1. The active share  $|Active_k|$  is instrumented by two indicator variables corresponding to stocks switching between Russell 1000 and 2000 in either direction. We use the estimates for  $\mathcal{E}_{agg,k}$  from the model with a variable value of  $\chi$  over time. Columns 2 and 3 show the corresponding OLS and reduced-form regressions, respectively. Column 4 is the first-stage regression. Standard errors are clustered by date.

## H Appendix Figures

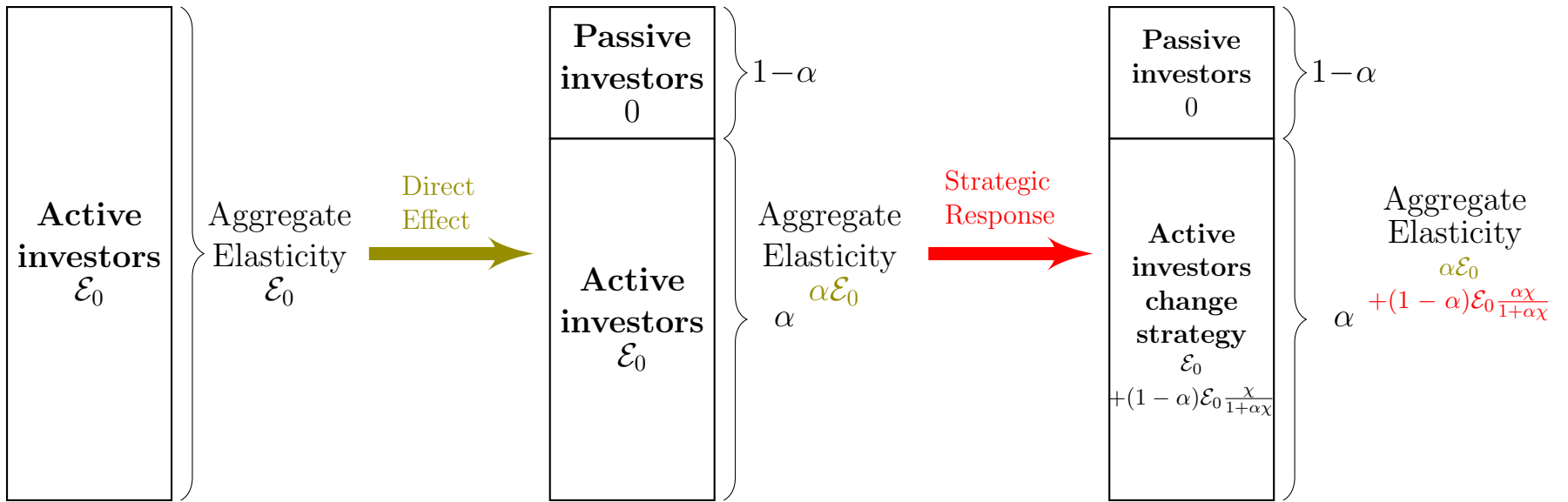
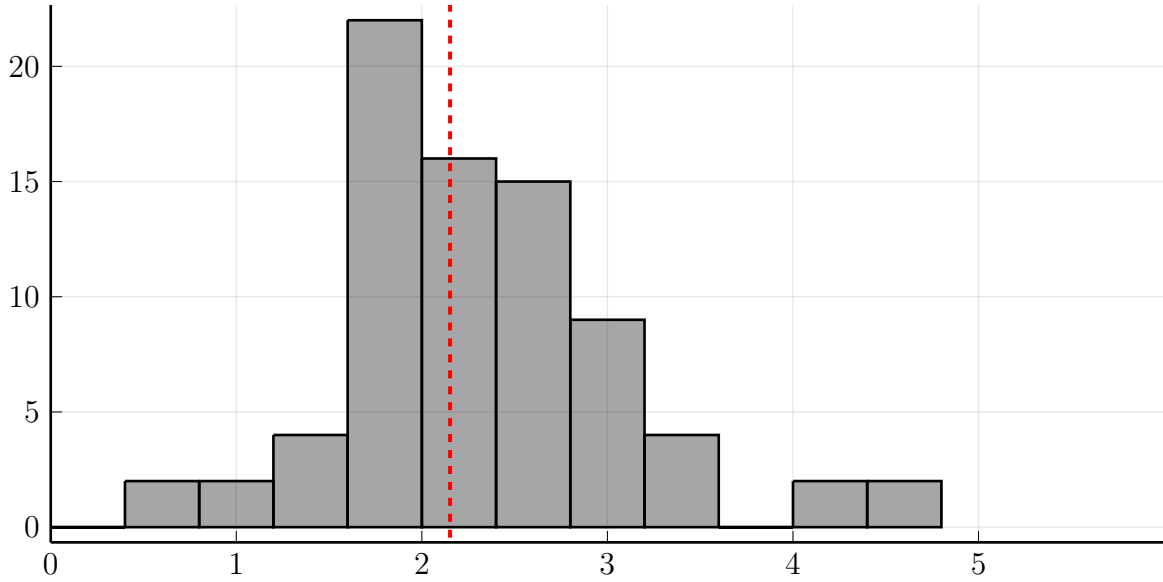
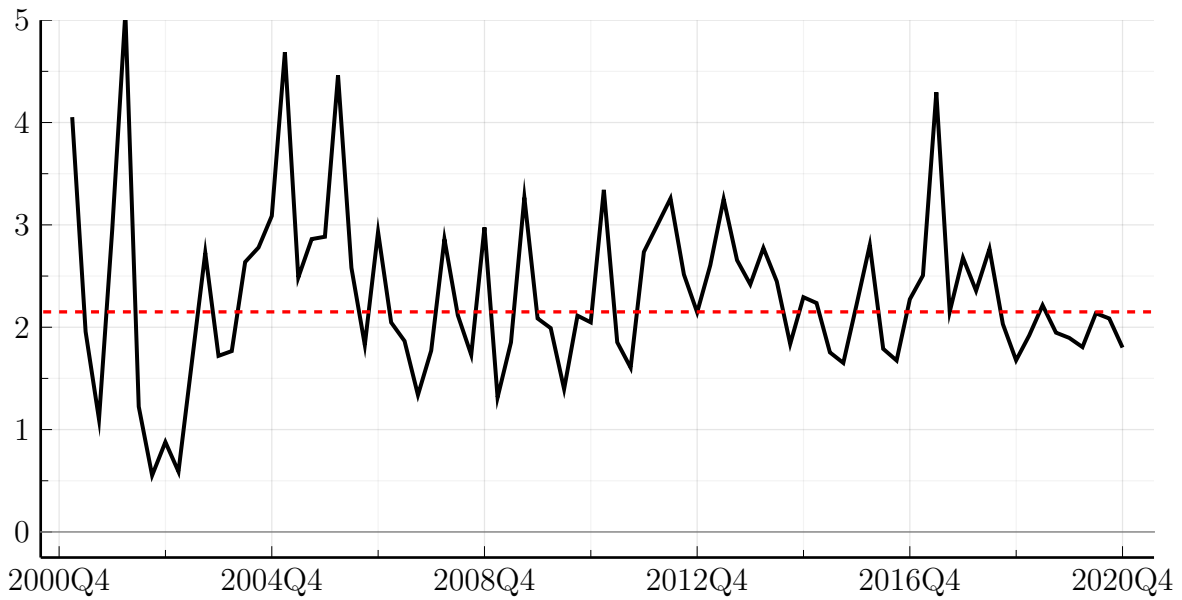


Figure IA.2. Effect of an increase in passive investing.

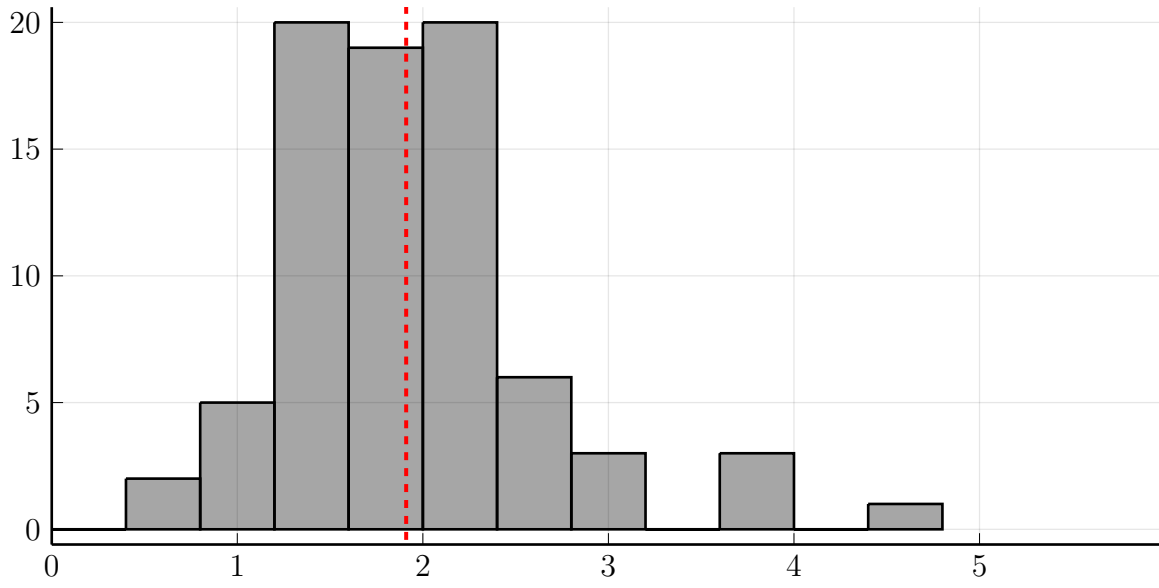


**Figure IA.3. Distribution of the degree of strategic response  $\chi$  across dates.** Figure IA.3 presents an histogram of our estimates of the degree of strategic response  $\chi$  for each date between 2001 and 2020. The median estimate over the time-period is  $\chi = 2.15$  (dashed red line).

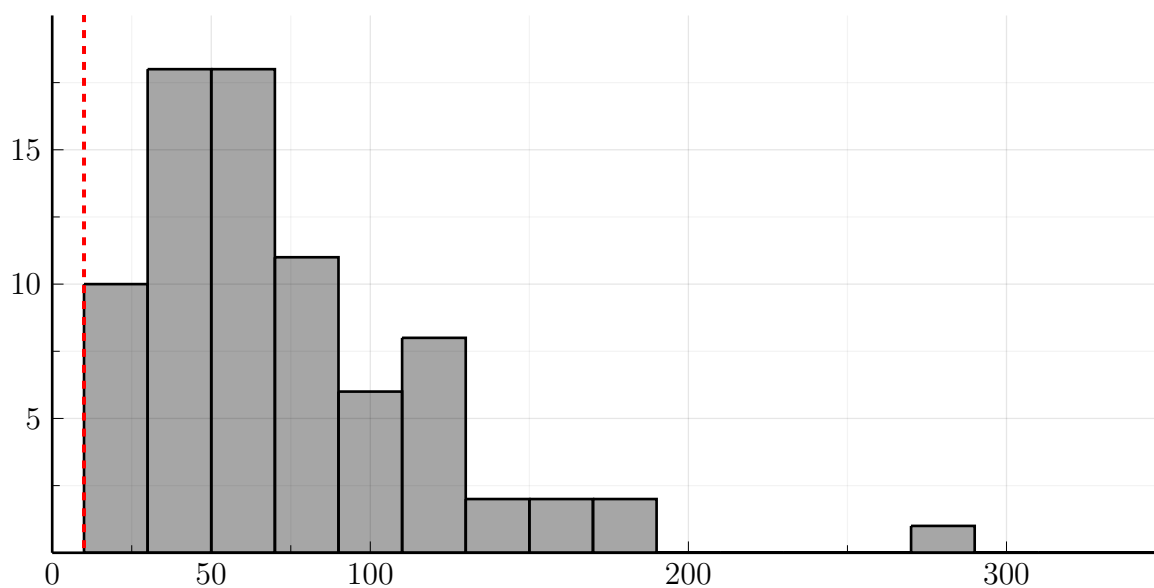


**Figure IA.4. Time-series of the degree of strategic response  $\chi$ .** Figure IA.4 shows the time-series of the estimates for the degree of strategic response  $\chi$ . The median estimate over the time-period is  $\chi = 2.15$  (dashed red line).

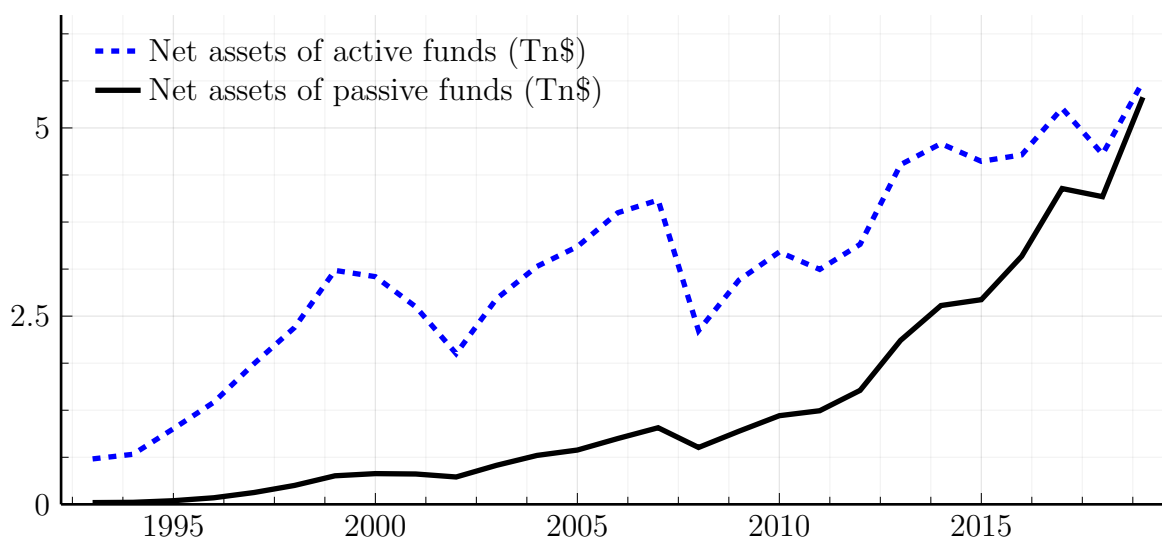




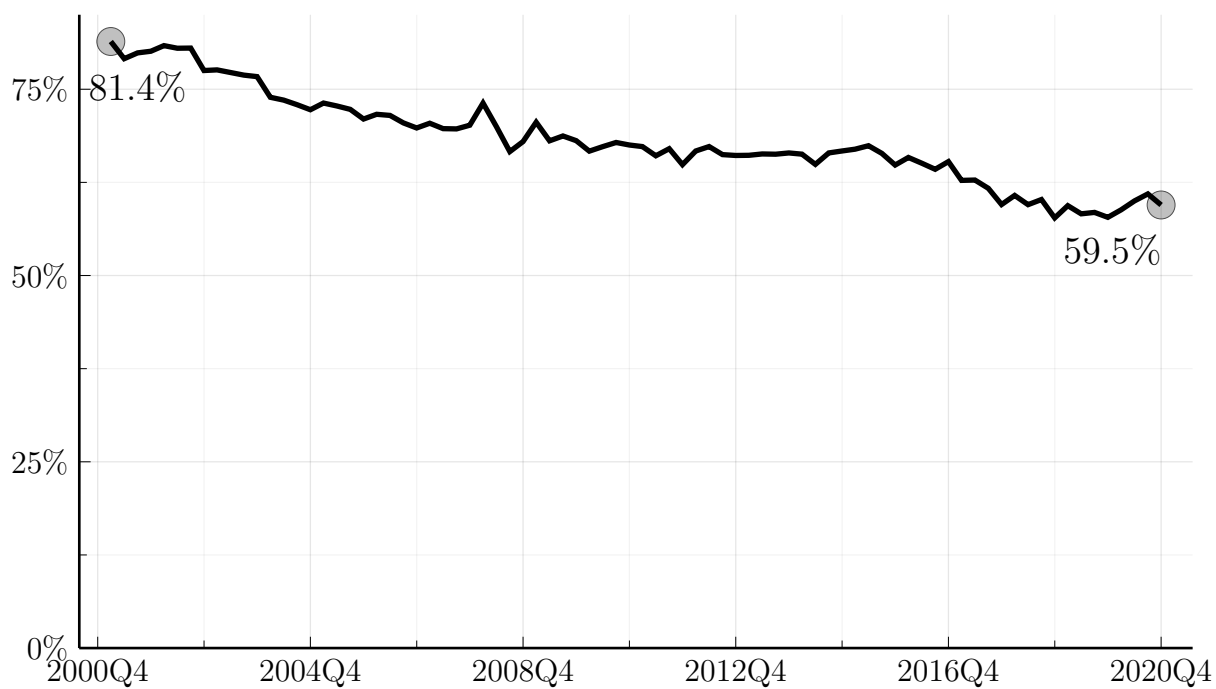
**Figure IA.5. Estimates of  $\chi$  using the book-equity weighted instrument.** Figure IA.5 presents an histogram of our estimates of the degree of strategic response  $\chi$  where the instrument for aggregate elasticity weights portfolios by book equity, for each date between 2001 and 2020. The median estimate over the time-period is  $\chi = 1.91$  (dashed red line).



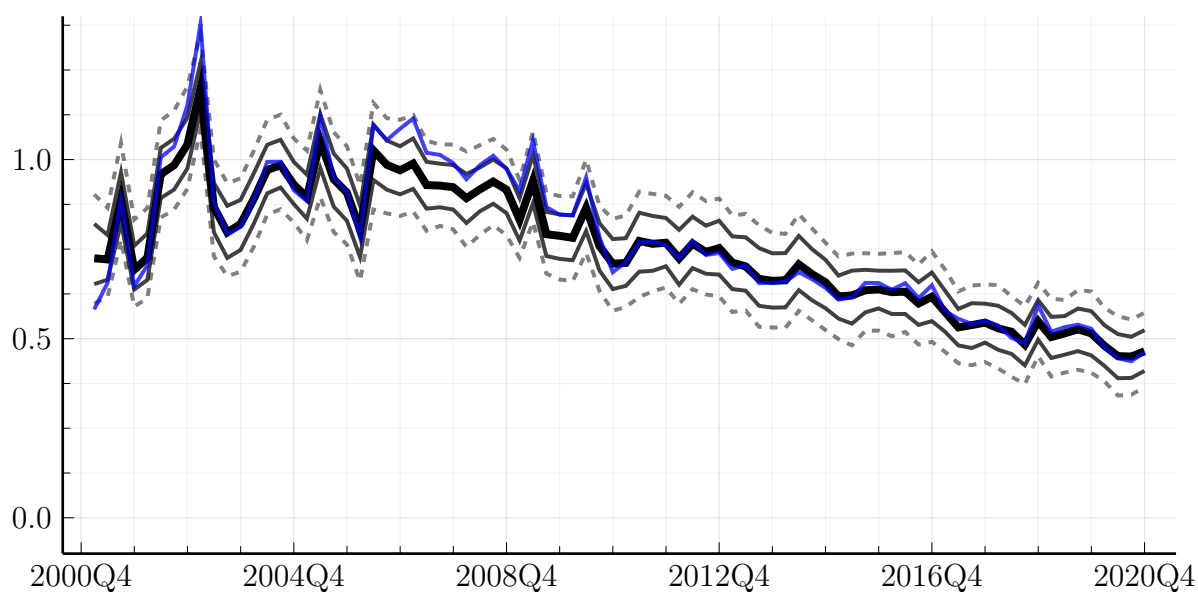
**Figure IA.6. Relevance condition for the the book-equity weighted elasticity instrument.** Figure IA.6 shows the first-stage F-statistics (Kleibergen-Paap) for the instrument of price times the aggregate elasticity when the instrument of the aggregate elasticity is weighted by book equity. The dashed red line is the critical value of 10.



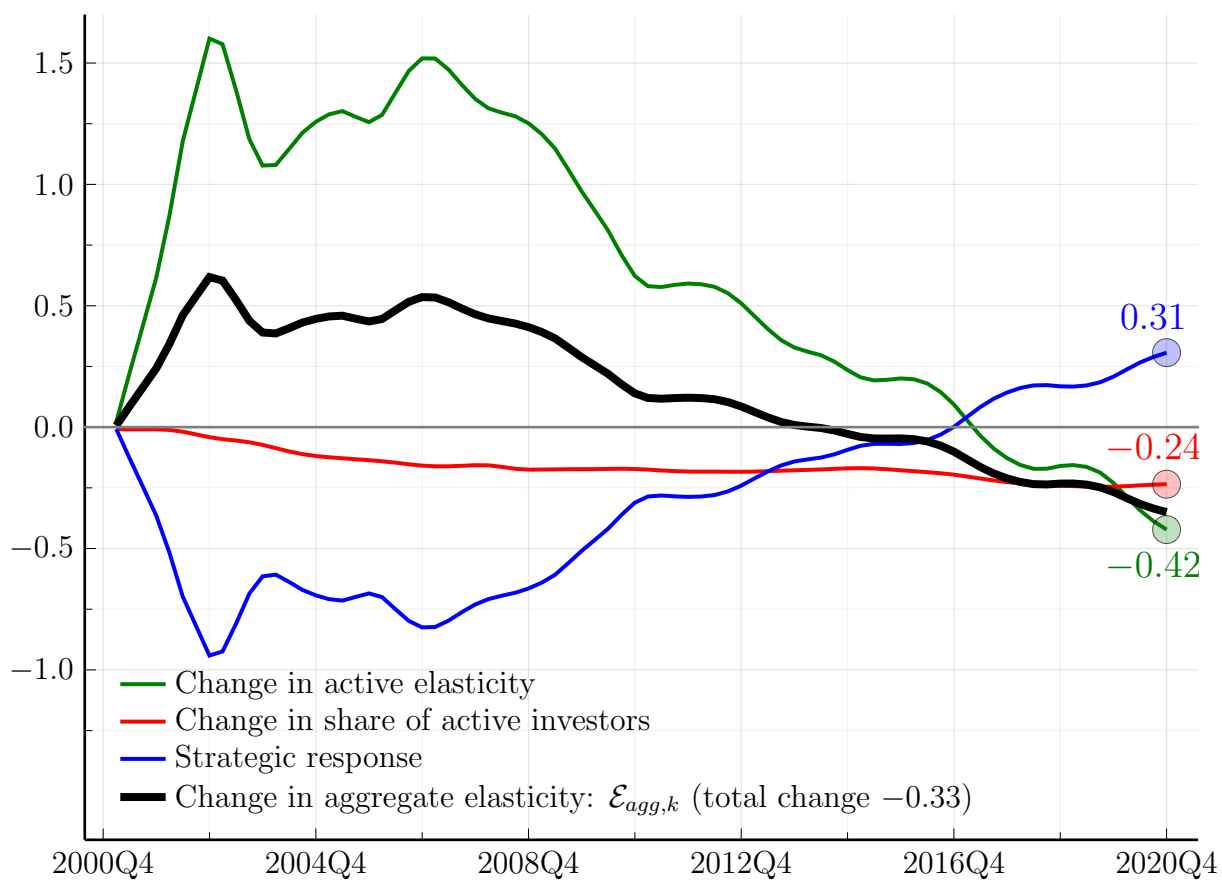
**Figure IA.7. Net assets of passive and active funds.** Figure IA.7 shows the net assets of domestic mutual funds and ETFs in trillions of dollars (year-end) for passive funds (black solid line) and active funds (blue dashed line). Source: ICI (2020).



**Figure IA.8. Fraction of active investors.** Figure IA.8 reports the fraction of active investors according to our model. For each stock, we compute the ratio of total position of active investors and the market capitalization. We report the median across stocks.



**Figure IA.9. Distribution of individual-specific elasticities  $\underline{\mathcal{E}}_{ik}$ .** Figure IA.9 shows the quantiles of the distribution of individual elasticities  $\underline{\mathcal{E}}_{ik}$  across investors for each stock and each date. We average the quantiles for each date to plot their time series. The black bold line is the average across investors. The two thin grey lines represent the 25th and 75th percentiles. The two dashed grey lines represent the 10th and 90th percentiles. And the solid blue line represents the average individual elasticities of the household investor.



**Figure IA.10. Decomposition of the change in aggregate elasticity for the the book-equity weighted elasticity instrument.** Figure IA.10 shows the decomposition derived in equation (28) over time, based on elasticities estimated from the book-equity weighted instrument for  $\mathcal{E}_{agg,k}$ . We compute each term of the decomposition for each date and accumulate the changes over time.

## References

2020. *Investment Company Fact Book – A Review of Trends and Activities in the Investment Company Industry*. Investment Company Institute, 60 ed.
- Backus, Matthew, Christopher Conlon, and Michael Sinkinson. 2019. “Common ownership in America: 1980-2017.” Tech. rep., National Bureau of Economic Research.
- Backus, Matthew, Christopher T Conlon, and Michael Sinkinson. 2020. “Common Ownership Data: Scraped SEC form 13F filings for 1999-2017.”
- Brunnermeier, Markus K. and Lasse Heje Pedersen. 2008. “Market Liquidity and Funding Liquidity.” *The Review of Financial Studies* 22 (6):2201–2238.
- De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. “Noise Trader Risk in Financial Markets.” *Journal of Political Economy* 98 (4):703–738.
- Duffie, Darrell. 2010. “Presidential Address: Asset Price Dynamics with Slow-Moving Capital.” *The Journal of Finance* 65 (4):1237–1267.
- Duffie, Darrell and Bruno Strulovici. 2012. “Capital Mobility and Asset Pricing.” *Econometrica* 80 (6):2469–2509.
- Fama, Eugene F. and Kenneth R. French. 1993. “Common risk factors in the returns on stocks and bonds.” *Journal of Financial Economics* 33 (1):3–56.
- . 2015. “A five-factor asset pricing model.” *Journal of Financial Economics* 116 (1):1–22.
- Haddad, Valentin, Paul Ho, and Erik Loualiche. 2021. “Bubbles and the Value of Innovation.”
- Koijen, Ralph S. J. and Motohiro Yogo. 2019. “A Demand System Approach to Asset Pricing.” *Journal of Political Economy* 127 (4):1475–1515.
- Kyle, Albert S. 1989. “Informed Speculation with Imperfect Competition.” *The Review of Economic Studies* 56 (3):317–355.
- Miller, Edward M. 1977. “Risk, Uncertainty, and Divergence of Opinion.” *Journal of Finance* 32 (4):1151–1168.
- Mitchell, Mark and Todd Pulvino. 2012. “Arbitrage crashes and the speed of capital.” *Journal of Financial Economics* 104 (3):469–490. Market Institutions, Financial Market Risks and Financial Crisis.
- Mitchell, Mark, Todd Pulvino, and Erik Stafford. 2002. “Limited Arbitrage in Equity Markets.” *The Journal of Finance* 57 (2):551–584.
- Shleifer, Andrei and Lawrence H. Summers. 1990. “The Noise Trader Approach to Finance.” *Journal of Economic Perspectives* 4 (2):19–33.

Shleifer, Andrei and Robert W. Vishny. 1997. “The Limits of Arbitrage.” *The Journal of Finance* 52 (1):35–55.

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan. 2012. “The short of it: Investor sentiment and anomalies.” *Journal of Financial Economics* 104 (2):288–302. Special Issue on Investor Sentiment.

———. 2015. “Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle.” *The Journal of Finance* 70 (5):1903–1948.

Veldkamp, Laura L. 2011. *Information Choice in Macroeconomics and Finance*. Princeton University Press.