The Long-Run Evolution of the Financial Sector

Maryam Farboodi* and Laura Veldkamp†

February 12, 2016‡

Abstract

As the financial sector has swollen in size, the nature of its activities has shifted and investors have taken ever larger bets on its outcomes. The financial sector should add value by processing information, evaluating risks, disseminating that information through advising or road shows, and ultimately, using the information to allocate capital. Over time, the nature of the information processed and transmitted has changed. While financial analysis used to mean fundamental analysis of an asset's long-run value, perhaps combined with some statistical exploration of recent price trends, more recently, focus has shifted to mining order flow data to identify promising times at which to trade. We build a model that explores the reason for this analysis shift, offers testable predictions to help determine the extent of the shift, and clarifies the consequences for the real economy.

As the financial sector has swollen in size, the nature of its activities has shifted and investors have taken ever larger bets on its outcomes. Why? Much of what the financial sector does is process information, evaluate risks, disseminate that information through advising or road shows, and ultimately, use the information to make investment decisions. Over time, the nature of the information processed and transmitted has changed. While financial analysis used to mean fundamental analysis of an asset's long-run value, perhaps combined with some statistical exploration of recent price trends, more recently, focus has shifted to mining order flow data to identify promising times at which to trade (see Hendershott, Jones, and Menkveld (2011)). While circumstantial evidence and many anecdotes suggest that the nature of financial analysis has changed, measuring analysis directly is nearly impossible. We cannot see what information analysts, fund managers or investment advisors acquire. Yet, understanding the nature of the information this industry is producing and the incentives for that production is crucial for policy makers and regulators deciding what to prohibit, to disclose or to tax. We develop theory that predicts whether changes in the economic environment should shift analysis, produces a set of testable predictions, and clarifies the economic consequences.

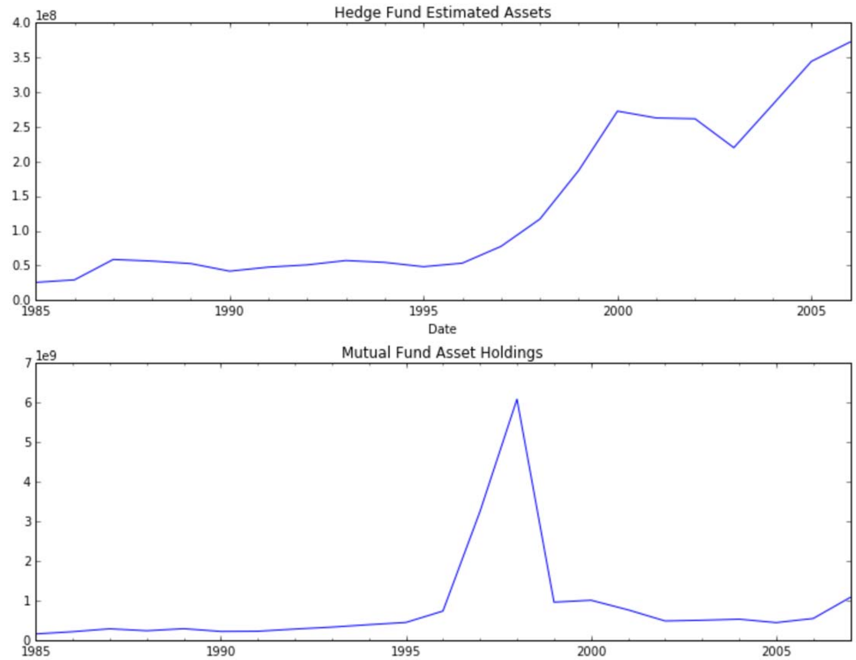
To explore a shift in financial analysis, we need a new type of information choice model, one that allows investors to choose between styles of financial analysis and then observe the information from that analysis, before they invest. This model reveals that the combination of growth in household wealth and the financial value of listed firms can explain shifts in the financial sector's size, leverage and analysis style.

*Princeton University; farboodi@princeton.edu

†This draft is highly incomplete and should not be circulated externally or cited without permission of the authors. Department of Economics Stern School of Business, NBER, and CEPR, New York University, 44 W. 4th Street, New York, NY 10012; lveldkam@stern.nyu.edu; http://www.stern.nyu.edu/∼lveldkam.

‡We thank Markus Brunnermeier, Martin Eichenbaum and Sergio Rebelo for comments and Goldman Sachs for their financial support through the GMI Fellowship program. JEL codes: Keywords: financial intermediation, liquidity, information acquisition.

Figure 1: Assets under management: Hedge funds are growing more quickly than mutual funds.
Sources: WRDS, the TASS data, series s12 and s34. The data is monthly and spans January 1985-January 2007. We average that data by year to avoid missing values and smooth data error. Units are $100 million for hedge funds and $billion for mutual funds.



This framework then allows us to explore the consequences of these trends for real output, investor welfare and the balance sheet exposure of the most-informed financial institutions to unexpected shocks.

The driving force behind the model is the endogenous change in the value of fundamental versus order-flow information. When information is scarce, it is very valuable to know what the fundamental value of an asset is. But when most investors are well-informed, it becomes more valuable to identify and trade against the remaining non-informational trades. Order flow analysis allows investors to target these more profitable trades.

Suggestive evidence of the trend we are exploring comes from the change in the mix of hedge funds and mutual funds. While mutual funds have traditionally done research of firm characteristics, "kicked the tires" by digging into firm data and investigating operations, and analyzed earnings statements, many hedge funds employ different modes of analysis. Many of their trading strategies center around the use of order-flow data. They often partner up with a broker who alerts them to the timing of uninformed trades, allowing them to take the other side of those trades. Figure 1 illustrates the explosion of hedge funds, relative to mutual funds. While this data surely cannot pinpoint fund analysis styles. It is suggestive of the type of transformation that financial intermediation has witnessed.

Contribution to the existing literature Glode, Green, and Lowery (2012) model growth in fundamental analysis but has no real spillover that could potentially make financial analysis valuable. Our theory allows for a choice between styles, which is important because more fundamental analysis will always be bad for investor welfare (Hirschliefer) and good for output. The shift from fundamental to technical analysis has more subtle effects.

Philippon (2015) argues that increased issuance can explain the growth of the financial sector because of intermediation per share has been relatively stable. This framework allows us to look inside the black box of intermediation and ask what information intermediaries collect and how it affects welfare and the real economy. The idea that there is growth in information processing is supported by the rise in price informativeness documented by Bai, Philippon, and Savov (2013).

In the information frictions literature in both finance and macro, papers such as Goldstein, Ozdenoren, and Yuan (2013), Bond, Edmans, and Goldstein (2012) and Angeletos, Lorenzoni, and Pavan (2010) have explored different ways in which asset prices spill over to real economic investment or efficiency. Our economic spillover, based on the effort a manager exerts when his compensation is based on the stock price, is slightly different from these mechanisms. Yet, we do not claim it is a better mechanism. It is simply a tractable way to capture the idea that there are real efficiency benefits to having an informationally efficient stock price. These existing papers all offer richer theories of how financial and real economies interact.

In the microstructure literature, Hendershott and Menkveld (2014) and Hendershott, Jones, and Menkveld (2011) document trends in trading activity. They use natural experiments to ask how this change from fundamental to algorithmic trading affects liquidity. Without a model, these papers cannot explain why the change took place. Similarly, empirical work alone would have a hard time identifying real effects of the long-run shift in analysis. A quantitative model is helpful here to gauge how large these effects might be.

A small, growing literature examines how the incentives to acquire order-flow information depend on others' information choices. In Yang and Ganguli (2009), agents can choose whether or not to purchase a fixed bundle of fundamental and order-flow information. In Yang and Zhu (2016), traders assigned to trade first get fundamental information while second-round traders choose order-flow information. But neither of these papers examines the choice that is central to this paper: The choices of whether to process more about asset payoffs or to analyze more order flow. Without that trade-off, these papers cannot explore how that mix of chosen information changes over time as the economy evolves.

1 Model

To explore the evolution of financial analysis style and its consequences, we set up as simple a model as possible that has a choice of types of information to analyze, a portfolio choice, realistic wealth effects, and real spillovers from this asset market.

We begin by describing the firm's investment problem. The idea this part of the model captures is that there is real value in having an informationally efficient equity market because it induces more efficient real investments. There are many ways one might model this connection. Since the heart of our idea lie in the asset market, we model this real spillover as simply as possible.

Entrepreneur's Real Investment Decision Time is discrete and infinite. There is a single firm whose output y_t depends on a firm manager's labor choice l_t. Specifically, $y_t = g(l_t)$, where g is increasing and concave. Because effort is unobserved, the manager's pay w_t is tied to the equity price p_t of the firm: $w_t = \bar{w} + p_t$. However, effort is costly. We normalize the units of effort so that a unit of effort corresponds to a unit of utility cost. Insider trading laws prevent the manager from participating in the equity market. Thus the manager's objective is

$$U_m(l_t) = \bar{w} + p_t - l_t \tag{1}$$

The firm produces $y_t = g(l_t) + \epsilon_{yt}$ units of output at the end of the period, where $\epsilon_{yt} \sim N(0, \tau_0^{-1})$. This output is the payoff of the equity claim on the firm. The key friction here and the reason that asset price informativeness will promote efficiency is that the entrepreneur's investment choice is unobserved by equity investors. Thus, the entrepreneur only has an incentive to invest to the extent that price reflect and respond to the true investment.

Investor preferences and endowments At each date t, a measure-one continuum of infinitely-lived investors is born. Investors have CRRA utility for end of period-t consumption c_{it}:

$$U(c_{it}) = \frac{C_{it}^{1-\gamma}}{1-\gamma} - \chi_i \quad \text{for } \gamma > 1 \tag{2}$$

where χ_i is a utility cost of processing information.

Each investor i is endowed with a non-tradeable tree that yields e_{it} units of consumption goods in each period t. Investors can eat this endowment directly, or invest it in a tradeable risky asset. The risky asset is a claim to a share of the production y_t of the entrepreneur's firm. Because we don't want to layer an unnecessary consumption-savings problem on top of our model, we assume that firm output and endowments completely depreciate at the end of each period. If investor i chooses to purchase q_{it} shares of the risky asset at price p_t, i's end of period t consumption will be

$$c_{it} = e_{it} - p_t q_{it} + y_t q_{it} \tag{3}$$

The value of endowments is correlated with the firm's output: $e_{it} = \bar{e} + h_{it} y_t + \tilde{e}_{it}$, where $\bar{e}$ is known and $\tilde{e}_{it} \sim N(0, \tau_e^{-1})$ is independent across agents and independent of all the other shocks in the economy. The variable h_{it} governs the correlation of agent i's endowment with output. That variable has a common component and an investor-specific component: $h_{it} = x_t + \epsilon_{hit}$ where $x_t \sim N(0, \tau_x^{-1})$ and $\epsilon_{hit} \sim N(0, \tau_h^{-1})$.

The reason for this rich, correlated endowment process is that it preserves a motive to acquire information. For information to have value, prices must not perfectly aggregate asset payoff information. We inject noise in prices by giving investors both informational and non-information reasons for trade. They have non-financial income that they want to hedge with financial assets. Shocks to this non-financial income will create shocks to their hedging demand, which is our source of noise in prices.

Information Choice If we want to examine how the nature of financial analysis has changed over time, we need to have at least two types of analysis to choose between. Financial analysis in this model means signal acquisition. This acquisition could represent the cost of researching and uncovering new information. But it could also represent the cost of interpreting and computing optimal trades based on information that is readily available from public sources.

Investors choose how much information to acquire or process about each of two random variables: They can choose how much to learn about f, the payoff or fundamental value of the asset and also choose how much to learn about x, the hedgers' demand shocks, which is the source of non-fundamental fluctuations in prices. We call $\eta_{fit} = y_t + \epsilon_{fit}$ a fundamental signal and $\eta_{xit} = x_t + \epsilon_{xit}$ an order-flow signal. What investors are choosing is the precision of these signals. In other words, if the signal errors are distributed $\epsilon_{fi} \sim N(0, \Omega_{fi}^{-1})$ and $\epsilon_{xi} \sim N(0, \tilde{\Omega_{xi}}^{-1})$, then the precisions Ω_{fi} and $\tilde{\Omega_{xi}}$ are choice variables for investor i. For notational convenience, we define $\Omega_{xi} = \tau_h + \tilde{\Omega_{xi}}$. Instead of choosing $\tilde{\Omega_{xi}} \geq 0$, we then allow the investor choose $\Omega_{xi} \geq \tau_h$. Then Ω_{xi} represents the joint signal precision that the investor has both from order-flow analysis and from observing his own endowment exposure to systemic financial risk.

The utility cost of processing information takes the form $\chi_i = \chi_f \Omega_f^2 + \chi_x(\Omega_x - \tau_h)^2$. For each type of information, the cost is a simple quadratic. This represents the idea that getting more and more precise information about a given risk is tougher and tougher. But acquiring information about multiple risks is just linear. An investor could hire another equal size staff to perform the other kind of analysis with the same precision at the same cost.

Information sets and equilibrium An equilibrium is a sequence of effort choices by managers $\{l_t\}$ and information choices $\{\Omega_{fit}\}$, $\{\Omega_{xit}\}$ and portfolio choices $\{q_t\}$ by investors such that

1. Entrepreneur's choice i_t maximizes $E[u_m(l_t)]$ in (1), taking as given prior knowledge of the environment that implies a set of investor information choices and an equilibrium pricing rule.

2. Given the knowledge of prior distributions of variables only, and without knowing the entrepreneur's decision i_t, investors choose signal precisions Ω_{fi} and Ω_{xi} to maximize $E[ln(E[U(c_{it})|\eta_{fit}, \eta_{xit}, h_{it}, p_t])]$, where U is defined in (2), subject to $\Omega_{fi} \geq 0$ and $\Omega_{xi} \geq 0$.

3. Given the information set $\{k_{t-1}, \eta_{fit}, \eta_{xit}, p, h_{it}\}$, and taking asset prices and the actions of other agents as given, investors choose their risky asset investment q_{it} to maximize $E[U(c_{it})|\eta_{fit}, \eta_{xit}, h_{it}, p_t]$, subject to the budget constraint (3).

2 Analytical Results

2.1 Optimal portfolio choice

To solve our CRRA portfolio problem, we use a log-linear approximation. Since $\gamma > 1$, utility is negative. Therefore, we linearize $-ln(-U(C_{it}))$ by setting it equal to $\rho_{it}C_{it}$. If ρ_{it} can be a function of C_{it}, we could choose ρ so as to equate $\rho_{it} = [(\gamma-1)ln(C_{it}) + ln(\gamma-1)]/C_{it}$, in which case, the two utility functions would be identical. But then $\rho_{it}C_{it}$ would not be a linear function of C_{it}. Instead, we allow ρ_{it} to be a function of $E_t[C_{it}]$, where t denotes the beginning of period t information set, prior to any information processing. Thus, we approximate $-ln(-U(C_{it})) \approx [(\gamma-1)ln(E_t[C_{it}]) + ln(\gamma-1)]/E_t[C_{it}] \cdot C_{it}$. This log-linear approximation implies that utility is

$$U(C_{it}) \approx -exp\left[-\left((\gamma-1)ln(E_t[C_{it}]) + ln(\gamma-1)\right)\frac{C_{it}}{E_t[C_{it}]}\right]$$

We can then rewrite this log-linear approximation in a form that is like exponential utility $-\exp(-\rho_{it}C_{it})$, with a coefficient of absolute risk aversion

$$\rho_{it} \equiv [(\gamma-1)ln(E_t[C_{it}]) + ln(\gamma-1)]/E_t[C_{it}]. \tag{4}$$

This form of risk aversion introduces wealth effects on portfolio choices.

Each investor chooses a number of shares q of the risky asset to maximize (2.1), subject to the budget constraint (3). The first-order condition of that problem is

$$q_{it} = \frac{E[f_{it}|\mathcal{I}_{it}] - p_t}{\rho_{it} Var[f_{it}|\mathcal{I}_{it}]} - h_{it} \tag{5}$$

Given this optimal investment choice, we can impose market clearing (**??**) and obtain a price function that is linear in asset payoffs and hedging shocks:

$$p = A + Bf + Cx \tag{6}$$

where A is in the appendix and the coefficients B and C are the solution to the following set of equations:

$$B = 1 - \tau_0/\bar{\Omega} \tag{7}$$

$$C = -\left(\frac{\bar{\Omega} - \tau_0}{2\bar{\Omega}\Omega_f}\right)\left[\bar{\rho} + \sqrt{\bar{\rho}^2 - 4\Omega_f\Omega_x}\right] \tag{8}$$

where $\bar{\Omega}$ (solved for in eq. 27) is the average precision of investors' posterior beliefs about y_t and $\bar{\rho} \equiv (\int 1/\rho_i di)^{-1}$ is the harmonic mean of investors' risk aversions.

2.2 Optimal firm investment

The firm manager chooses his effort to maximize (1). The first order condition is $Bf'(l_t) = 1$, which yields an equilibrium effort level $l_t = (f')^{-1}(1/B)$. Notice that the socially optimal level would set the marginal utility cost of effort (1) equal to the marginal product $f'(l)$. Instead the manager sets this marginal product equal to $1/B$. When B is below one, managers under-provide effort, relative to the social optimum because their stock compensation moves less than one-to-one with the true value of their firm.

2.3 Analysis Style Choice

After we substitute the optimal portfolio choice (5) and the equilibrium price rule (6) into utility (2.1) and take the expected value, we obtain an indirect utility function, expressed in terms of signal precision of fundamental information Ω_f and order-flow information Ω_x :

$$\max_{\Omega_{fi},\Omega_{xi}} \lambda\Omega_{fi} + \phi\Omega_{xi} - \chi_f\Omega_{fi}^2 - (\chi_x - \tau_h)^2. \tag{9}$$
$$\text{s.t. } \Omega_{fi} \geq 0, \text{ and } \Omega_{xi} \geq \tau_h.$$

The first order conditions are

$$\Omega_{fi} = \frac{\lambda}{2\chi_f} \tag{10}$$

$$\Omega_{xi} = \tau_h + \frac{\phi}{2\chi_x} \tag{11}$$

Note that since $\phi = (B/C)^2\lambda$, then if the two information costs are equal ($\chi_f = \chi_x$) ratio between the amount of fundamental and the amount of technical information acquired is always $(B/C)^2$. For now, we assume that the two information costs are equal, until we relax this assumption at the end. Note also that because of the quadratic cost functions, as long as λ and ϕ are positive, the non-negativity constraints never bind.

2.4 Household Wealth Effects on Financial Analysis

Wealth changes absolute risk aversion ρ. Show how this changes the amount and the composition of financial analysis.

Proposition 2.1 *An increase in wealth makes the marginal value of order-flow information ϕ rise by a greater proportion than the marginal value of fundamental information λ.*

Figure 2: Evolution of financial analysis and asset prices as wealth rises.
Parameter values are $\gamma = 5$, $\tau_0 = 1$, $\bar{x} = 1$, $\tau_x = 0.1$, $\tau_h = 0.5$, $\chi_f = 20$, and $\chi_x = 1$.

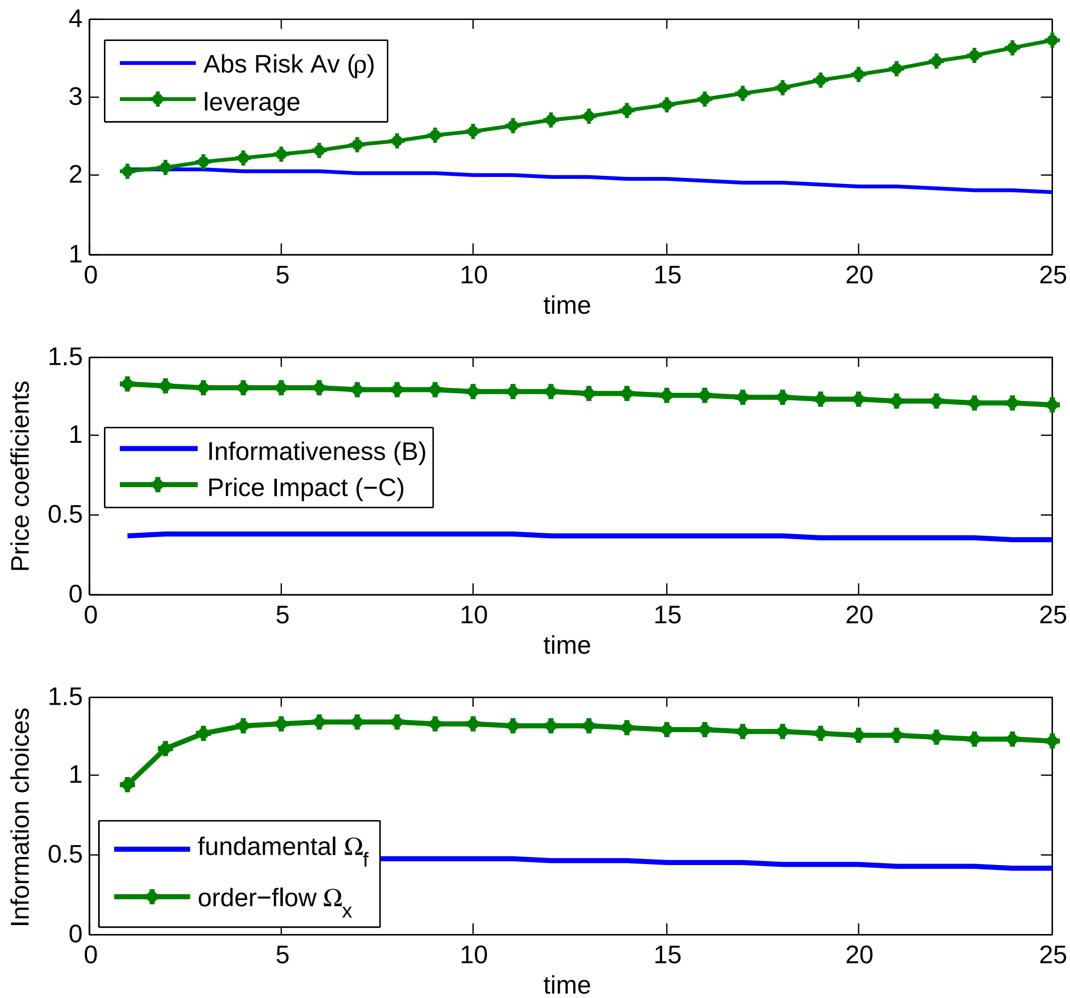


Figure 2 illustrates these effects. It plots the equilibrium price coefficients, information choices, and leverage in an economy where wealth starts at 2 and grows at 2.5% per year, corresponding to an average rate of income growth in the U.S.. The top panel shows that as wealth rises, leverage rises. Agents who are less risk averse take larger positions in risky assets. Without endogenous information, the risky asset position $q = 1/\rho\bar{\Omega}_i E[y_t - p_t|\mathcal{I}]$ would just scale up as ρ falls and $1/\rho$ rises. Since ρ falls at rate $ln(W)/W$, the risky asset position would grow more slowly than wealth, and the difference $W - q * p$ can fall. With information acquisition, the difference rises as investors borrow more and more to take larger bets on risky assets.

The middle panel of Figure 2 shows that changes in risk aversion have little impact on either price informativeness or the price impact of an uninformed trade. But the rise in wealth can explain the explosion of order-flow analysis, as shown in the bottom panel.

2.5 Firm Growth Effects on Financial Analysis

This is an increase in asset supply. The firm investment problem will tell us what parameter of the firm problem needs to grow to get them to issue more shares.

Proposition 2.2 *Holding other investors' choices fixed, a marginal increase in firm shares $\bar{x}$ increases fundamental and order-flow analysis proportionately.*

In equilibrium, an increase in firm shares $\bar{x}$ increases the amount of order-flow analysis by a larger proportion, if and only if $\rho^2 - 4\Omega_f\Omega_x > 4\Omega_x^2$.

Figure 3 illustrates the effects of a rise in asset supply. It plots the equilibrium price coefficients, information choices, and leverage in an economy where shares start at 1 and grow at 10% per year. The top panel shows that as the size of the market grows, leverage falls. This is because the increase in asset supply reduces the price of shares.

The middle panel of Figure 3 shows that a growth in shares causes a slight rise in price informativeness and a large drop in the price impact of an uninformed trade. The a rise in the number of shares can also explain the explosion of order-flow analysis, as shown in the bottom panel.

2.6 Highly-Informed Financial Institutions

What if some investors have different wealth (?) or information costs (?). How will the leverage of those agents be different? Leverage here is their position in riskless assets, divided by initial wealth. Can we show that these agents are more susceptible to some kind of surprise shocks (wealth, payoff, hedging)?

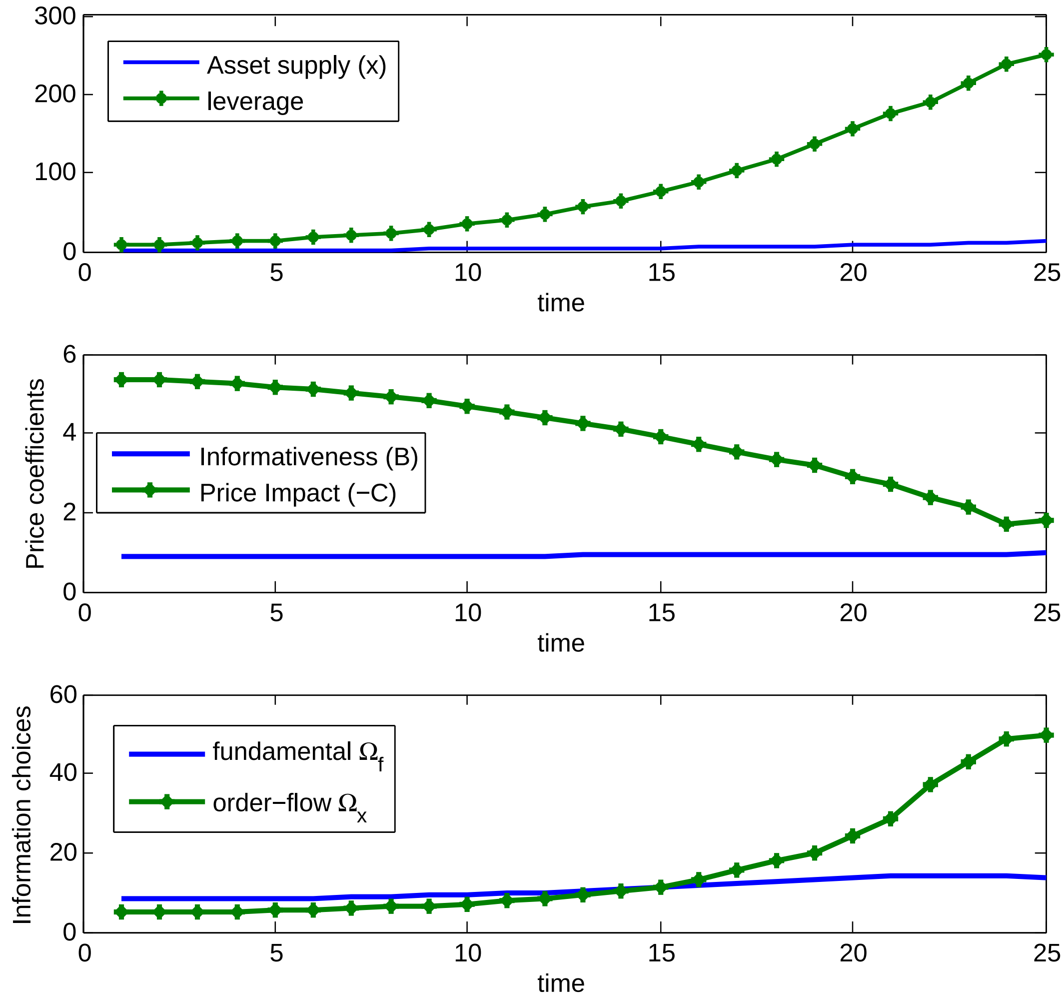
2.7 An Uncertain Future

Show that as analysis grows, the linear equilibrium disappears. This could mean a market crash. It could mean a jump in price as some non-linear equilibrium emerges. Perhaps these are the increasingly frequent flash crashes we see. But it is clear that when both fundamental and order-flow analysis are sufficiently precise, the laws governing market dynamics will change.

3 Calibration

The asset-market parameters that we need to match to data include γ (CRRA risk av), $E[f]$ (or equiv marg cost of firm investment), $V[f]$,$\bar{x}$, τ_x, χ, α. Our calibration strategy follows that of Kacperczyk, Nosal,

Figure 3: Evolution of financial analysis and asset prices as the number of shares grows.
Parameter values are $\gamma = 5$, $\tau_0 = 1$, $\tau_x = 0.1$, $\tau_h = 0.5$, $\chi_f = 20$, and $\chi_x = 1$. Asset supply $\bar{x}$ starts at 1 and grows exponentially at 10% per year.



and Stevens (2015).

Calibration targets:

1. We need a time series of W to map CRRA into CARA. HH leverage? Exclude housing wealth? Look at Hyun and Adrian (liquidity and leverage paper).
   Household wealth data? Include pension funds and institutional investors too? Or is this part of HH wealth?

2. $E[f]$ doesn't matter. We could normalize it. Only determines a price level. As long as it doesn't change over time. But if price informativeness changes and firms issue more valuable assets, does that change the problem? No, still not. All these expected values will be reflected in price and won't affect asset holdings or information.

3. prior variance: unconditional variance of market return? or market price?

4. Asset supply with market cap? Divide the market cap of every firm by its price. This gives shares

per firm. Add these shares up. Compute quarterly to get a time series. (Not a balanced panel - we want new firms to add to shares.)

5. information processing costs: Make two costs equal and match the price informativeness at the start of the sample? At the start of our sample, regress p_t on $p_{t+1} + d_t$ and get A, B and C. Use chi and τ_x to match B and $C^2\tau_x^{-1}$ (variance of residual).

   An alternative strategy: Calibrate the information processing cost by matching the portfolio return differences of sophisticated and retail investors. KNS say this difference is: 14% vs 11.3% for 1989-2000.

6. Let's choose CRRA parameter to match A. Falling CARA will cause returns to fall. But a growth in firm shares will offset this.

7. α Call this the fraction of shares held by non-institutional investors.
   From Kacperczyk, Nosal, and Stevens (2015) 23% sophisticated in 1989-2000 and 43% sophisticated for 2001-2012. Can we calibrate to 23% and get a trend in the right direction? This is a fraction of shares held.

4 parameters need to match endogenous objects in the model (Sim Meth Moments): A, B C and α.

4 Numerical Results

Use the calibrated model to measure the effects on outcomes of interest. These are potentially interesting targets for measurement. We don't need to include them all.

4.1 Evolution of Analysis Styles

The story would be that there has always been some order-flow analysis done, primarily by market makers or specialists. But there wasn't widespread demand for this kind of information because there wasn't lots of fundamental information in prices. Once the fundamental analysis sector grew, the demand for order-flow information took off.

When wealth is low, investors start by doing mostly fundamental analysis. As wealth grows, risk aversion falls, information processing increases and shifts to processing more order-flow information.

Include price informativeness in here.

Compare this to hedge fund data. How the covariance of fund holdings with macro or price noise shocks changes over time.

4.2 Size of the financial sector

This is resources spent on all information processing. Show how much it increases. How much can wealth alone explain? What if we also include the growth in shares (firm size)? Even if we get only a fraction of the growth, that tells us something.

4.3 Leverage

As wealth rises, investors take larger gambles. They also acquire more information, which encourages them to take even more extreme positions, based on the information they learn. These large long or short positions often involve negative riskless positions, which we interpret as leverage. The following results show how leverage growth with wealth and firm growth.

4.4 Real investment

Liquidity matters here.

4.5 Exposure of Institutions to Unexpected Shocks

Institutions would be agents who process more information. Because their conditional variance is lower, they should take larger bets and be more levered. How does this tendency change over time? Could we call this systemic risk? How does it change the sensitivity of their asset value to shocks like an extreme negative payoff shock, an uncertainty shock, or an increase in risk aversion (decline in non-equity wealth). Ideally, can we measure the growth of systemic risk?

4.6 Investor utility

4.7 Trends in Equity Premium

Does the increase in analysis cause the equity premium to fall?

5 Conclusion

References

Angeletos, G.-M., G. Lorenzoni, and A. Pavan (2010): "Beauty Contests and Irrational Exuberance: A Neoclassical Approach," NBER Working Paper.

Bai, J., T. Philippon, and A. Savov (2013): "Have Financial Markets Become More Informative?," NBER Working Paper 19728.

Bond, P., A. Edmans, and I. Goldstein (2012): "The Real Effects of Financial Markets," *Annual Review of Financial Economics*, 4(1), 339–60.

Glode, V., R. Green, and R. Lowery (2012): "Financial Expertise as an Arms Race," *Journal of Finance*, 67, 1723–1759.

Goldstein, I., E. Ozdenoren, and K. Yuan (2013): "Trading frenzies and their impact on real investment," *Journal of Financial Economics*, 109(2), 566–82.

Hendershott, T., C. Jones, and A. Menkveld (2011): "Does Algorithmic Trading Improve Liquidity?," *Journal of Finance*, 66, 1–34.

Hendershott, T., and A. Menkveld (2014): "Price Pressures," *Journal of Financial Economics*, 114, 405–423.

Kacperczyk, M., J. Nosal, and L. Stevens (2015): "Investor Sophistication and Capital Income Inequality," Imperial College Working Paper.

Philippon, T. (2015): "Has the U.S. Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation," *American Economic Review*, 105(4), 14081438.

Yang, L., and J. Ganguli (2009): "Complementarities, Multiplicity, and Supply Information," *Journal of the European Economic Association*, 7(1), 90–115.

Yang, L., and H. Zhu (2016): "Back-Running: Seeking and Hiding Fundamental Information in Order Flows," .

A Proofs

A.1 Bayesian Updating

To form the conditional expectation, $E[f_{it}|\mathcal{I}_{it}]$, we need to use Bayes' law. But first, we need to know what signal investors extract from price, given their observed endowment exposure h_t and their order-flow signal η_x. For an investor who learned nothing about order flow ($E[x]=0$) the information contained in prices is $(p-A)/B$ from the linear price equation (6), we see that this is equal to $y+C/Bx$. Since x is a mean-zero random variable, this is an unbiased signal of the asset payoff f.

But conditional on h_t and η_x, x_t is typically not a mean-zero random variable. Instead, investors use Bayes' law to combine their prior that $x_t=0$, with precision τ_x with their endowment and order flow signals: h_{it} with precision τ_h and η_{xit} with precision Ω_{xi}. The posterior mean and variance are

$$E[x|h_{it},\eta_{xit}] = \frac{\tau_h h_{it} + (\Omega_{xi}-\tau_h)\eta_{xit}}{\tau_x+\Omega_{xi}} \tag{12}$$

$$V[x|h_{it},\eta_{xit}] = \frac{1}{\tau_x+\Omega_{xi}} \tag{13}$$

Then, the unbiased price signal, conditional on order flow info is $\eta_{pi} \equiv (p-A-C\,E[x|\eta_x])/B$. Since that is equal to $f+C/B(x-E[x|\eta_x])$, the variance of price signal noise is $(C/B)^2Var[x|\eta_x]$. In other words, the precision of the price signal for agent i is $\tau_{pi}\equiv(B/C)^2(\tau_x+\Omega_{xi})$.

Now, we can use Bayes' law for normal variables again to form beliefs about the asset payoff. We combine the prior μ, the price/order-flow information η_{pi}, and the fundamental signal η_{fi} into a posterior mean and variance:

$$E[y_t|\mathcal{I}_{it}] = (\tau_0+\tau_{pi}+\Omega_{fi})^{-1}\left(\tau_0\mu+\tau_{pi}\eta_{pi}+\Omega_{fi}\eta_{fi}\right) \tag{14}$$

$$V[y_t|\mathcal{I}_{it}] = (\tau_0+\tau_{pi}+\Omega_{fi})^{-1} \tag{15}$$

Average expectations and precisions: Next, we integrate over investors i to get the average conditional expectations. Begin by considering average price information. The average price informativeness is $\tau_p\equiv(B/C)^2(\tau_x+\Omega_x)$. The average fundamental and order flow signal precisions, averaged over all investors are $\Omega_f\equiv\int\Omega_{xi}di$ and $\Omega_x\equiv\Omega_{xi}di$. This allows us to form the average posterior precision:

$$\bar{\Omega}\equiv\int V[y_t|\mathcal{I}_{it}]^{-1}di = \tau_0+(B/C)^2(\tau_x+\Omega_x)+\Omega_f. \tag{16}$$

Since the signal realizations are uncorrelated with their precisions (more precise signals are not positive or negative on average), we can integrate the weights and signals separately to get the the average conditional expectation-variance ratios. For the hedging shock x, we can use the formulas for η_{pi} and τ_{pi}, cancel terms and integrate to get $\int\tau_{pi}\eta_{pi}di$ $=\tau_p(p-A)/B-(B/C)\Omega_x x_t$. Then we can use this formula to express the conditional expectation-variance ratio of f as a linear combination of the prior, price and x_t info and the firm value y_t:

$$\int\frac{E[y_t|\mathcal{I}_{it}]}{Var[y_t|\mathcal{I}_{it}]}di = \tau_0\mu_t+\tau_p\frac{p-A}{B}-\frac{B}{C}\Omega_x x_t+\Omega_f y_t. \tag{17}$$

A.2 Solving for equilibrium prices

To solve for equilibrium prices, start from the portfolio first-order condition for investors (5) and equate total demand with total supply:

$$\int\frac{E[f_{it}|\mathcal{I}_{it}]-p_t}{\rho_{it}Var[f_{it}|\mathcal{I}_{it}]}-h_{it}\;di=\bar{x} \tag{18}$$

Integrating over all agents[1] and using (17) we can rewrite

$$\frac{1}{\bar{\rho}}\left[\tau_0\mu_t+\tau_p\frac{p-A}{B}-\frac{B}{C}\Omega_x x_t+\Omega_f y_t\right]-\frac{1}{\bar{\rho}}\bar{\Omega}p_t-x_t=\bar{x} \tag{19}$$

[1]Keep in mind that risk aversion and information choice may be linked and thereby induce correlation in ρ_{it} and the conditional variance term

where $1/\bar{\rho} \equiv \int 1/\rho_{it} di$ is average risk tolerance ($\bar{\rho}$ itself is a harmonic mean) and recalling that $\int h_{it} di = x_t$. Rearranging this expression, we can write

$$p_t = \frac{1}{\bar{\Omega} - \tau_p/B}\left[\tau_0\mu_t - \frac{\tau_p A}{B} - \frac{B}{C}\Omega_x x_t + \Omega_f y_t - \bar{\rho}(\bar{x} + x_t)\right] \tag{20}$$

matching coefficients yields

$$A = \frac{\tau_0\mu_t - \bar{\rho}\bar{x}}{\bar{\Omega}} \tag{21}$$

and $B = 1 - \tau_0/\bar{\Omega}$, which is (7).

The formula for C is a quadratic equation. Coefficient matching yields $C = -(\bar{\rho} + \Omega_x B/C)/(\bar{\Omega} - \tau_p/B)$. Notice that C appears on the left and also on the right in the denominator. This produces the quadratic form. We can substitute in the solution for B and write out the quadratic form as

$$\bar{\Omega}C^2 + \frac{\bar{\Omega} - \tau_0}{\Omega_f}\bar{\rho}C + \frac{(\bar{\Omega} - \tau_0)^2}{\bar{\Omega}\Omega_f}\Omega_x = 0 \tag{22}$$

Applying the quadratic formula yields two solutions. We use a continuity argument to select one of the roots. When there is no information about order flow, $\Omega_x = 0$, the B/C term on the right side of the C expression disappears and $C = -\bar{\rho}(\bar{\Omega} - \tau_p/B)$. The smaller root of C converges to this value as $\Omega_x \to 0$. But the larger root converges to 0, making the B/C term in price informativeness go to ∞. Since we are interested in the evolution of the economy as order-flow analysis grows continuously, the root with the continuous limit is more helpful for our analysis. Therefore, we proceed by focusing on this solution, which is given in (8).

But even this is not a complete characterization of the solution because $\bar{\Omega}$ depends on the average information in prices τ_p, which in turn depends on B/C. Thus, the solution to the equilibrium pricing coefficients is a set of B, C and $\bar{\Omega}$ values that solve the three equations jointly.

A.3 Deriving information choice objective

If we substitute q_{it} into the budget constraint (3) and then into expected utility (**??**) and take the time-2 expectation, we get $EU_2 = -exp(-(1/2)Var[y_t - p_t|\mathcal{I}]^{-1}E[y_t - p_t|\mathcal{I}]^2)$. To get time-1 expected utility, we take the log of $-EU_2$, which simply cancels the exponential, and then take the negative prior expectation; $EU_1 = (1/2)E[Var[y_t - p_t|\mathcal{I}]^{-1}E[y_t - p_t|\mathcal{I}]^2]$. As of time-1, the conditional expectation $E[y_t - p_t|\mathcal{I}]$ is a random variable because the signal realizations it is conditioned on are not known. But it is a normally distributed random variable: $E[y_t - p_t|\mathcal{I}] \sim N((1-B)\mu - A,\ V_E)$. The variance $V_E = (1-B)^2/\tau_0 + CC'/\tau_x - Var[y_t - p_t|\mathcal{I}]$ is computed using the law of total variance: the ex-ante variance of a conditional expectation is the prior variance, minus the posterior variance. Since the conditional expectation enters squared, we use the formula for the mean of a non-central chi-square to compute expected utility: $EU1 = (1/2)Var[y_t - p_t|\mathcal{I}]^{-1}[((1-B)\mu - A)^2 + V_E]$.

Thus, the information choice problem becomes

$$\max_{\Omega_{fi},\Omega_{xi}} \lambda\Omega_{fi} + \phi\Omega_{xi} - \chi_f\Omega_{fi}^2 - \chi_x\Omega_{xi}^2 \tag{23}$$

where the following are the aggregate variables in the problem:

$$\phi = \left(\frac{B}{C}\right)^2\lambda \tag{24}$$

$$\lambda = (1-B)^2\tau_0^{-1} + C^2\tau_x^{-1} + (\rho\bar{x}\bar{\Omega}^{-1})^2 \tag{25}$$

From the equilibrium pricing equations (7) and (8), we can express

$$\frac{B}{C} = \frac{1}{2}\Omega_x^{-1}\left(-\rho + \sqrt{\rho^2 - 2\Omega_f\Omega_x}\right) \tag{26}$$

and recall that $B = 1 - \bar{\Omega}^{-1}\tau_0$ (from price solution, equation 7). From Bayes' law, we can express the average posterior precision as a sum of the prior precision, plus the average fundamental signal precision, plus the average precision of the information inferred jointly from order flow and prices:

$$\bar{\Omega} = \tau_0 + \Omega_f + \left(\frac{B}{C}\right)^2 (\tau_x + \Omega_x). \tag{27}$$

A.4 Comparative static wrt asset supply

A.5 Comparative static wrt absolute risk aversion

Proof: If we reduce the costs $\chi = \chi_f = \chi_x$, this affects Ω_{fi} and Ω_{xi} for all agents i through the first order condition. So, we then ask how B/C...

A.6 Comparative static wrt symmetric information cost

A.7 Comparative static wrt order-flow information cost

B Data

For now, let's try to record everything we look into here so that we have a repository of data sources we might use.

1. WRDS: TASS has 13F filings for hedge funds over $100mil in assets under management.