Information Production by Intermediaries: Relative Valuation and Balanced Designs

Armando Gomes
Washington University in St. Louis

Alan Moreira
University of Rochester

David Sovich
Washington University in St. Louis

17 Encontro Brasileiro de Financas

Motivation

- Financial analysts are often employed to produce information about the future value of risky assets, e.g.:
  - Equity analysts producing stock recommendations to investors
  - Employees evaluating investment projects or recruiting candidates for their employer
- Investor then aggregates analysts' reported information in order to make a more informed decision, e.g:
  - Invest more (less) capital on projects that have higher (lower) reported values
  - Hire candidates with good evaluations

Motivation

This paper incorporates several realistic features often present in real world information production environments:

1. Analysts produce information for only a small subset of assets:
    - Sell-side stock analysts cover on avg. 7 stocks often in the same industry but sometimes in two industries (Gomes, et al (2016))
    - Credit analysts working for credit rating agencies follow on avg. 10 firms (Fracassi, Petry, and Tate (2016))
    - GPs at PE and VC firms are responsible for on avg., respectively, 3 and 5 companies (Metrick and Yasuda (2010))

Table 1: **Sample Statistics - Firms and Analysts**

This table contains summary statistics for the analysts and the firms in our sample. Our sample of firms is comprised of all firms in the S&P 500 in 2012. Our sample of analysts is comprised of all analysts who cover one of the sample firms between $2012-2016$, where analyst coverage is defined as issuing an EPS estimate. This leaves us with 511 firms and 3,182 analysts in our sample. The variables in the Table are defined as follows: (1) **Firms Per Analyst** - the number of distinct sample firms that an analyst covers during the sample period, (2) **Industries Per Analyst** - the number of distinct GICS Industries that an analyst covers during the sample period, (3) **Analyst Per Firm** - the number of distinct analysts that a firm is covered by during the sample period, and (4) **Distinct Connections Per Firm** - the number of stocks that a firm is connected to (as defined by sharing at least one common analyst with) in the sample.

Dependent Variable	N **(1)**	Mean **(2)**	StDev **(3)**	P10 **(4)**	P25 **(5)**	P50 **(6)**	P75 **(7)**	P90 **(8)**
Firms Per Analyst	3,182	5.54	4.37	1	2	4	8	12
Industries Per Analyst	3,182	1.83	1.21	1	1	1	2	3
Analysts Per Firm	511	34.46	14.49	18	24	33	43	54
Distinct Connections Per Firm	511	51.80	17.17	32	40	50	62	74

Source: “Analyst Coverage Network and Corporate Financial Policies”, by Gomes, Gopalan, Leary, and Marcet (2016)

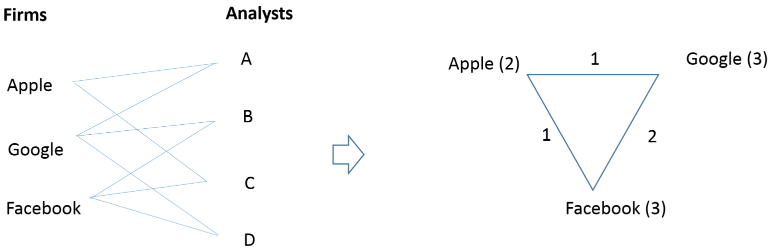
Motivation

2. Analysts are better at producing relative valuation rather than absolute asset valuations
   - Stock analysts may know how to compare the value of Apple vs Google or Walmart vs Target but have no clue about their absolute value
3. Analysts produce information that contains an analyst-specific fixed-effect
   - An extensive literature in finance and accounting documents that stock recommendations and target prices of sell-side analysts are biased (Barber et al (2001))
   - Fracassi, Petry, and Tate (2016) find that credit analysts exhibit systematic optimism and pessimism affecting debt prices and credit spreads

Research Questions

- What is the optimal allocation of information producing resources in the presence of these realistic features?
- How should investors efficiently aggregate all the information produced by various analysts to decide how much capital to invest in each asset?

Information Production Network



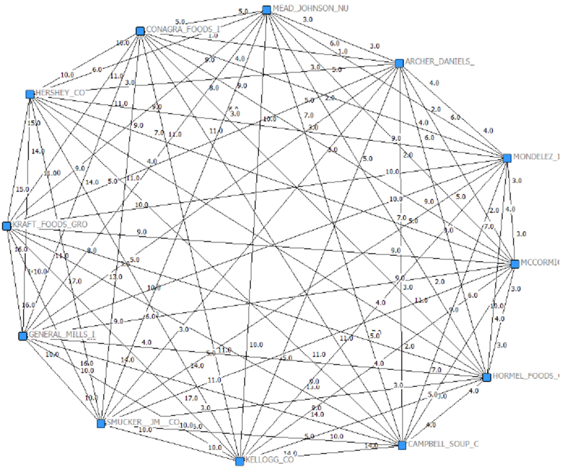
Information Production Network

- The information producion network is the graph where the vertices are the firms/assets and the edges are the number of common analysts weighed by the precison of the link
  - Connection to graph theory helps understand economic properties

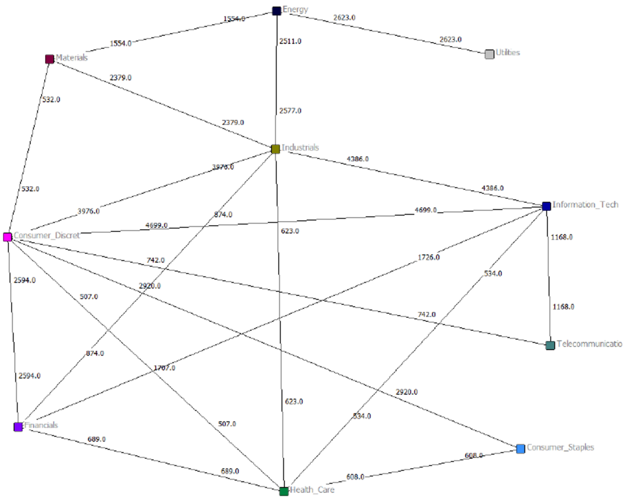
Summary of Main Results

- The information production network plays key role
- Under relative valuation capital is only reallocated between the connected components of the information network
- The intensity of the capital reallocations depends on both the value of the relative signals and the strength of the connections between the assets
  - Empirically, analysts covering S&P 500 stocks primarily follow stocks in the same industry, making connections among industry stocks strong
    - Thus investors can create value by reallocating capital among stocks with relatively high vs low recommendations (Boni and Womack (2006))

Common Analysts among S&P 500 Food Companies



Common Analysts among Industries/Sectors



Summary of Main Results

- The optimal information production network displays a balancedness property in which pairs of distinct assets are followed by approximately equal number of analysts
  - Intuitively, this allows for more efficient use of the relative valuation between assets
- Investor utility under relative valuation is enhanced by broad asset coverage with analysts spreading their time evenly across assets covered
- Analysts optimally learn about assets with both extremely high and low excess returns in order to create long-short portfolios

Related Literature

- Information acquisition and asset prices
  - Stijn and Veldkamp (2010), Kacperczyk, Stijn, and Veldkamp (2016), Mondria (2010)
- Rational Inattention
  - Sims (JME 2003), Peng and Xiong (2006), Peng (2005)
- Optimal experimental block design
  - Kiefer (1975), Silvey (1980), and modern surveys by Pukelsheim (SIAM 2006) and Chaloner and erdinelli (1995)

Model

- Risk-averse investor with CARA or mean-variance preference investing in n risky assets and a riskless asset
- The investor hires m analysts to produce noisy information about the asset values
- Timeline
  - t=1 (*Information Production Decision):* Matching of analysts to assets. Each analyst can cover at most $q \leq n$ assets
    - Analysts report information to the investor
  - t=2 *(Investment Decision):* Investor then decide how much to invest in the assets
  - t=3: Uncertain asset payoffs are realized

Information Production Decision

- The asset return $R = [R_1, ..., R_n]'$ has normal distribution $R \sim N(\bar{R}, \Sigma)$.
- Each analyst $a = 1, ..., m$ produces a signal y_{ia} about $i's$ asset return

$$y_{ia} = R_i + u_a + \varepsilon_{ia},$$

  - The analyst-specific noise term u_a has precision $\phi_a \geq 0$ where $u_a \sim N\left(0, \phi_a^{-1}\right)$
    - pure relative valuation case: $\phi_a = 0$;
    - pure absolute valuation case: $\phi_a = \infty$
  - The total analyst precision is $\tau_a > 0$ and $\theta_{ia} \geq 0$ is the fraction of time spent by analyst a on asset i, and $\varepsilon_{ia} \sim N\left(0, (\tau_a \theta_{ia})^{-1}\right)$
  - Each analyst a can produce information for at most q_a assets at convex or linear cost $c\left(\phi_a, \tau_a\right)$

Investment Decision

- In period 2, the investor observes the asset recommendations y and make his investment decision
- The investor allocates a fraction $\omega = [\omega_1, ..., \omega_n]'$ of his wealth to the risky assets and the remainder to the risk-free asset yielding return r_f
  - $\mu = \left(\bar{R} - r_f\right)$ is the excess return
- The investor total return R_p is equal to $R_p = \omega' \left(R - r_f\right) + r_f$

Investment Decision

- The investor has CARA or mean-variance preference where the expectations is taken over the joint distribution of returns and signals:

$$U = -\frac{1}{\gamma} \log \left[E \left[\exp \left(-\gamma R_p \right) \right] \right] \text{ CARA certainty equivalent utility}$$

$$U = E \left[E \left[R_p | y \right] - \frac{\gamma}{2} Var \left[R_p | y \right] \right] \text{ Mean-variance preference}$$

- For both preferences the optimal portfolio weights, given that analysts report signal y, is:

$$\omega = \frac{1}{\gamma} \left(\text{var} \left(R | y \right) \right)^{-1} \left(E \left(R | y \right) - r_f \right)$$

Investment Decision

- **Lemma:** Each analyst produce information matrix

$$\Theta_a = \tau_a \left(diag(\theta_a) - \frac{\tau_a}{\tau_a + \phi_a} \theta_a \theta_a' \right)$$

and the investor posterior mean and variance of return is:

$$E(R|y) = \left(\Sigma^{-1} + \sum_{a=1}^{m} \Theta_a \right)^{-1} \left(\sum_{a=1}^{m} \Theta_a y_a \right) + \bar{R},$$

$$var(R|y) = \left(\Sigma^{-1} + \sum_{a=1}^{m} \Theta_a \right)^{-1}$$

and the total precision/information matrix is $\Theta = \sum_{a=1}^{m} \Theta_a$

Information Matrix

- The information matrix Θ can be decomposed as

$$\Theta = \sum_{a=1}^{m} \left[\frac{\phi_a}{\tau_a + \phi_a} \underbrace{[diag(\tau_a \theta_a)]}_{\text{absolute valuation}} + \frac{\tau_a}{\tau_a + \phi_a} \underbrace{[\tau_a \left(diag(\theta_a) - \theta_a \theta_a' \right)]}_{\text{relative valuation}} \right]$$

- The first term is a diagonal matrix with the precision added by the absolute valuation part of the information
- The second term captures the relative valuation component

Investment Matrix

- The relative valuation information matrix is the *Laplacian matrix* of the information production network graph G:
  - The graph where the vertices are the firms and the edges are all the pairs of distinct firms that are covered by at least one common analyst
  - The Laplacian contains the relevant information about the strength of the connections among firms

Analyst Coverage Network

- Define the *weighted adjacency matrix* $\mathsf{A}(G)$ of the analyst coverage network as follows:
  - For any two adjacent firms $i \sim j$ connected by a common analyst, then
$$\mathsf{A}(G)_{ij} = \sum_a \tau_a \theta_{ia} \theta_{ja}.$$
- $\mathsf{A}(G)_{ij}$ is a measure of the relative strength of the connection between firms i and j.
- Define the *Laplacian matrix* $L(G)$ as
$$\mathsf{L}(G) := \mathsf{D}(G) - \mathsf{A}(G) = \sum_a \tau_a \left(diag(\theta_a) - \theta_a \theta_a' \right),$$
  where $\mathsf{D}(G)$ is the diagonal matrix with the degree $\sum_a \tau_a \theta_{ia}$ of firm i.

Wealth Reallocation

Proposition (Portfolio Choice) Given an precision matrix Θ the the optimal investor portfolio choice is

$$\omega(y) = \frac{1}{\gamma}\left(\sum_{a=1}^{m} \Theta_a y_a + \Theta\mu + \Sigma^{-1}\mu\right).$$

- In response to a signal received by analyst a for asset i, the optimal investment on assets $j \neq i$, and i changes by

$$\frac{\partial \omega_j}{\partial y_{ia}} = -\frac{1}{\gamma}\frac{\tau_a}{\tau_a + \phi_a}\theta_{ia}\theta_{ja} \leq 0 \text{ and } \frac{\partial \omega_i}{\partial y_{ia}} \geq 0$$

Wealth Reallocation

- **Proposition** (cont'd) In the pure relative valuation case, where $\phi_a = 0$:
  - There is no reallocation of wealth among disconnected components: i.e., for all signals y, $\sum_{i \in \mathcal{G}} \omega_i(y)$ is constant.
  - Moreover, there is no reallocation of wealth between risky and riskless assets, i.e., for all signals y, $\sum_{i \in N} \omega_i(y)$ is the same as under no information learning

Investor Welfare

Proposition: The ex-ante investor utility associated with an information production matrix Θ is
(i) In the CARA case the ex-ante utility is

$$U(\Theta) = \frac{1}{2\gamma}\left(\log\left(\det\left(I + \Sigma\Theta\right)\right) + \mu'\Sigma^{-1}\mu\right).$$

(ii) In the mean-variance case the ex-ante utility is

$$U(\Theta) = \frac{1}{2\gamma}\left(Tr\left(\Sigma\Theta\right) + \mu'\Theta\mu + \mu'\Sigma^{-1}\mu\right).$$

Investor Welfare

- **Lemma:** For any non-singular variance matrix Σ:
  (i) The investor utility function $\mathcal{U}(\Theta)$ of a CARA investor is **strictly monotonic** and **strictly concave** in the information matrix Θ.
  (ii) The investor utility function $\mathcal{U}(\Theta)$ of a mean-variance investor is **strictly monotonic** and **linear** in the information matrix Θ.
  - Therefore, the problem of optimally allocating information producing agents across assets is a concave maximization problem:

$$\max_{\Theta} \mathcal{U}(\Theta) - c(\Theta) \ \text{ s.t. } \Theta \text{ is feasible}$$

Relative versus Absolute Valuation

- **Proposition:** Suppose an analyst produces information about q stocks with attention θ and precision ϕ and τ, so that the precision matrix is $\Theta = \tau\left(diag(\theta) - \frac{\tau}{\tau+\phi}\theta\theta'\right)$. Then the incremental mean-variance utility gain, $\mathcal{U} = \frac{1}{2\gamma}\left(Tr\left(\Sigma\Theta\right) + \mu'\Theta\mu\right)$, can be expressed also as

$$\mathcal{U} = \frac{1}{2\gamma}\left[\begin{array}{c} \underbrace{\sum_{i,j}\tau\theta_i\theta_j\left[\left(var(r_i - r_j) + (\mu_i - \mu_j)^2\right)\right]\frac{\tau}{\phi+\tau}}_{\mathcal{U}_R} \\ \underbrace{+\tau\left(\sum_{i=1}^{q}\theta_i\left(var(r_i) + \mu_i^2\right)\right)}_{\mathcal{U}_A}\frac{\phi}{\phi+\tau} \end{array}\right]$$

Spreading Time Allocation evenly under Relative Valuation

- **Proposition:** Consider mean-variance investor with symmetric $\Sigma = \sigma^2 (I + \rho J)$. Then the optimal design is for analysts to spread their allocation of time equally across all q assets, i.e., $\theta_{ia} = \frac{1}{q}$, and to choose the precision ϕ^* and τ^* given by the unique global solution of the utility maximization problem

$$\mathcal{U} = \frac{\tau}{2\gamma} \left(\left(\sigma^2 (1 + \rho) + \mu^2 \right) \frac{\phi}{\tau + \phi} + \sigma^2 \left(\frac{q-1}{q} \right) \frac{\tau}{\tau + \phi} \right) - c(\tau, \phi)$$

  subject to $\tau, \phi \geq 0$.

- (Comparative statics) Increases in the correlation ρ and excess return μ, and decreases in q and volatility σ yields higher optimal precision ϕ^* and lower relative valuation precision τ^*.

Trade-off between Broad Coverage versus Specialization

- There are gains from broad coverage versus specialization under relative valuation as the utility gain is $\tau \times \left(\frac{q-1}{q}\right)$, but the marginal utility gain is decreasing with the asset coverage
  - There are substantial utility gains from increasing coverage from say 2 to 3 assets (approximately 33% gain) but the utility gain from increasing coverage from 9 to 10 assets is only approximately 1%
- However, the overall precision τ is likely decreasing in the total number of assets covered, $\tau'(q) < 0$, as agents may have limited processing capacity, thus the decision of how many firms to cover in a pure relative valuation setting involves solving $\max_q \tau\left(q\right) \times \left(\frac{q-1}{q}\right)$ which will lead to a relatively small number of assets being covered

Optimal Design with Asymmetric Assets: Long-Short Portfolios

- The optimal design with asymmetric assets and mean-variance preference where $\Sigma = diag(\sigma_1^2, ..., \sigma_n^2) + \sigma_f^2 J$, and the prior asset returns are $\mu_i = E\left(R_i\right) - r_f$
  - Under pure absolute valuation investors maximizes $\sum_{i=1}^{q} \theta_i \left(\sigma_i^2 + \sigma_f^2 + \mu_i^2\right)$ so only assets with maximum $\sigma_i^2 + \mu_i^2$ receive analyst attention.Under pure relative valuation maximizes $\sum_{i=1}^{q} \theta_i \left(\sigma_i^2 + \left(\mu_i - \bar{\mu}\right)^2\right)$ so analyst disperse their attention to a broader set of assets, and learns about assets with both extremely high and low excess returns in order to create **long-short portfolios.**

Optimal Design with Asymmetric Assets: Long-Short Portfolios

- **Proposition** (Asymmetric Assets - Mean-variance Utility) The unique global solution to the investor optimal design problem is for the analyst to allocate attention θ_i only to assets with values for

$$\sigma_i^2 + \frac{\phi}{\tau+\phi}\mu_i^2 + \frac{\tau}{\tau+\phi}\left(\mu_i - \bar{\mu}\right)^2 \geq \lambda$$

  above a certain cut-off value λ. The attention is given by

$$\theta_i = \frac{1}{2\sigma_i^2}\left(1+\frac{\phi}{\tau}\right)\left(\sigma_i^2 + \frac{\phi}{\tau+\phi}\mu_i^2 + \frac{\tau}{\tau+\phi}\left(\mu_i - \bar{\mu}\right)^2 - \lambda\right)^+$$

  where λ and $\bar{\mu}$ are constants obtained by the solution of the two equations $\sum_{i=1}^n \theta_i = 1$ and $\bar{\mu} = \sum_{i=1}^n \mu_i\theta_i$, and the function $x^+ := \max(x, 0)$.

Optimality of Balanced Designs

- The problem of optimally allocating information producing agents across assets when the investor has CARA utility is achieved with balanced allocations of information production resources.
  - **Definition:** A design is balanced if every pair of distinct assets (i, j) is covered by the same number of agents
- These allocations allow for the most efficient use of relative valuation information to extract the agent specific component of the signal.

Optimality of Balanced Designs

- **Example:** Consider $n = 6$, $m = 10$, and $q = 3$. Then the structure

$$\mathcal{A} = \{123, 124, 135, 146, 156, 236, 245, 256, 345, 346\}$$

denote the subset of assets followed by each of the $m = 10$ analysts.

  - The structure is a balanced design with each **pair** of assets covered by exactly 2 analysts and each asset is followed by exactly 5 analysts.

Optimality of Balanced Designs

Proposition: (Optimality of symmetric balanced designs) Consider the symmetric problem above where the investor has CARA preference with n assets and m agents that can cover $q \leq n$ assets with prior return $R \sim N\left(\bar{R}, \Sigma\right)$ where $\Sigma = \sigma^2 I + \sigma_f^2 J$. Then the most efficient allocation among all possible feasible allocations is the balanced design

Investor Welfare with a Balanced Design

Proposition : (Investors' utility under a balanced design) Consider the balanced design in which all m analysts choose the same precision τ and ϕ and each produce information about q assets and the agents are organized according to a balanced design with $\frac{mq(q-1)}{n(n-1)}$ analysts per pair of assets, and $\frac{mq}{n}$ analysts per asset.
(i) The precision of the signal obtained by investors is the $n \times n$ matrix Θ equal to

$$\Theta = \frac{\tau m}{n}\left[\frac{\phi}{\tau+\phi}I + \frac{\tau}{\tau+\phi}\frac{n\left(q-1\right)}{q\left(n-1\right)}\left(I - \frac{1}{n}J\right)\right],$$

where I and J are, respectively, the $n \times n$ identity matrix and matrix of ones in all entries

Investor Welfare with a Balanced Design

Proposition (cont'd) (ii) The expected investor utility gain with a balanced design is equal to:

$$\mathcal{U}(\tau,\phi) = \frac{1}{2\gamma}\left((n-1)\log(1+\lambda_1) + \log(1+\lambda_2)\right) - c(\tau,\phi)$$

where λ_1 and λ_2 are the eigenvalues of the weighted information matrix $\Theta\Sigma$ given by:

$$\lambda_1(\tau,\phi) = \sigma^2 \frac{m}{n}\frac{\tau}{\tau+\phi}\left(\phi + \tau\frac{n(q-1)}{q(n-1)}\right): \text{ (with multiplicity } n-1\text{)},$$

$$\lambda_2(\tau,\phi) = \sigma^2 \frac{m}{n}\frac{\tau\phi}{\tau+\phi}(1+n\rho): \text{ (with multiplicity } 1\text{)}$$

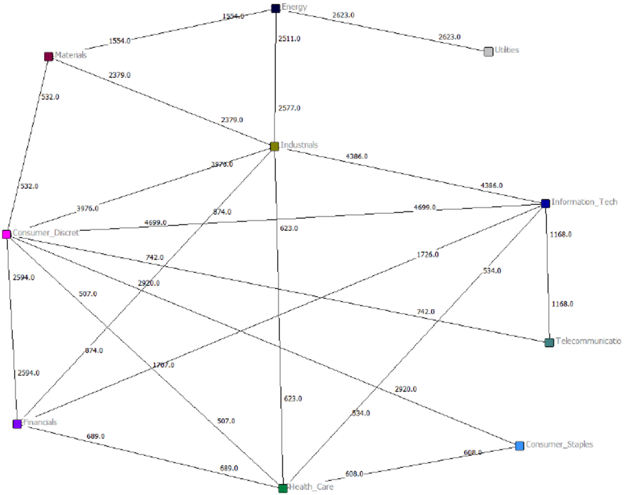
and the optimal precision τ^* and ϕ^* maximize the utility $\mathcal{U}(\tau,\phi)$ above

Information Production within and across Industries

Proposition: Consider a setting with investor with CARA utility and two industries each with n_j firms with variance $\Sigma = \Sigma_1 \oplus \Sigma_2$ with $\Sigma_j = \sigma_j^2 (I + \rho_j J)$

- Suppose there are a total of A analysts, each analyst can cover up to $q = 2$ assets, with relative precision τ_j for assets in industry j and τ_{12} for the cross-industry analysts
- Then the optimal allocation of analyst will have a mass of analysts x_j following industry j and a mass x_c of cross-industry analysts spread out in a block balanced pattern with any pair of assets in the same industry covered by the same number of analysts and any pair of assets from industry 1 and 2 also covered by the same number of analysts

Common Analysts among Industries/Sectors



Extensions

- In follow-up work we address the equilibrium information acquisition problem
  - In a multi-asset setting where the information produced by investors is partially reveled by prices (Admati (1982))
  - Where investors can produce realistic types of information such as relative, absolute, industry and market valuations
- Applications
  - What is the effet of indexing and the increased used of ETFs and passive vs active investing on the incentives to acquire information and on the informativeness of prices?
  - Our framework is very tracatable and allows for closed-form solutions which greatly help in applications

Conclusion

- Over 99% of all U.S. stocks belong to one giant connected component and the average and median distance between two stocks is less than three connections
- Connections are very strong among stocks belonging to the same industry, but there are also significant connections across-industry
- Our paper highlights the connectedness among U.S. stocks and the use of relative valuation and shows how to efficiently use the information stored in the analyst coverage network to invest in stocks