The Analyst Coverage Network

Armando Gomes, Alan Moreira, David Sovich

September 05, 2015

Abstract

We study the problem of aggregating biased analyst stock recommendations whose only informational content are relative valuations. In a Bayesian framework, we obtain closed-form expressions for the posterior mean and variance of excess returns and we derive the optimal portfolio policy for a utility maximizing mean-variance investor. We show that the optimal Bayesian portfolio is intricately linked to the analyst coverage network – 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 connectedness of the analyst coverage network determines how wealth is reallocated among components (e.g. industries) of stocks. Moreover, changes in portfolio depend not only on the value of relative stock recommendations, but also the strength of the connections between stocks in the network. Our model also admits a novel estimator for the consensus stock recommendation measure commonly used in the empirical evaluation of analyst forecasts.

How should investors incorporate sell-side analyst recommendations into their asset allocation decisions? While prior research demonstrates that sell-side recommendations carry investment value (Stickel (1995), Womack (1996), Barber et al. (2001,2006,2010), Jegadeesh et

*Armando and David are at Washington University at St. Louis. Alan is at Yale University. We thank seminar participants at Olin finance Brownbag for comments. Comments are welcome at gomes@wustl.edu.
al. (2004), Brav and Lehavy (2003)), it is a well-known empirical fact that these recommendations are biased (Lin and McNichols (1998), McNichols and O’Brien (1997)). Moreover, the sheer amount of stock recommendations produced by brokerages and financial institutions that is available to investors is astounding. In 2014 alone, there was a total of 3,970 sell-side analysts that produced a total of 32,554 recommendations and research reports. Investors are confronted then with the daunting task of how to properly incorporate an enormously large amount of informative, but potentially biased, signals into their investment decisions.

In this paper, we solve the problem of aggregating the information contained in a large network of analyst stock recommendations. Specifically, we solve the problem of an investor that is unable to evaluate the amount of bias contained in a specific analyst recommendation, and infer information only from relative signals. This information structure puts the analyst coverage network front and center. Intuitively, the investor needs the network to overlap in order for information about different stocks to be ordered. This implies that certain analysts, which cover stocks of very different industries, play a key role in enabling investors to learn and reallocated wealth across industries. This insight is empirically relevant because analyst coverage is strongly concentrated within industries, with 70% of stocks covered by the average analyst belonging to the same industry.

We propose a formal method for extracting information from a large cross-section of signals and solve the associated Bayesian portfolio choice problem. Using Bayes rule, we obtain closed-form expressions for the posterior mean and variance of excess returns and we derive the optimal portfolio policy for a utility maximizing mean-variance investor. The optimal portfolio is cleanly expressed as the sum of two components: how much investors would invest when receiving a zero posterior signal and how much they reallocate among stocks in response to stock recommendations. Our analysis nests the special case where analysts signals are unbiased and coincide with the optimal portfolio policy of Black and Litterman (1992).

To better understand the economics of the optimal Bayesian portfolio, we turn to the
study of the analyst coverage network - the graph that describes the extent to which firms are connected by common analysts. Interestingly, the matrix representation of this network plays a key role in understanding how investors reallocate their wealth in response to stock recommendations. First, the connectedness of the analyst coverage network determines how wealth is reallocated among groups (e.g. industries) of stocks. We show formally that there is no reallocation of wealth among disconnected components of the analyst coverage network. Intuitively, stock recommendations are only useful within connected components of the analyst coverage network and not between disconnected components because the presence of analysts biases makes investors are unable to compare information from disconnected components. Second, we can summarized the degree of capital reallocation across stocks using the Laplacian matrix and the Adjacency matrix of the coverage network. We show that reallocation of wealth depend not only on the value of relative stock recommendations, but also the strength of the connections between stocks in the network. Strong within-industry connections allow investors to make significant intra-industry changes in portfolio in response to stock recommendations. Additionally, for industries belonging to the same component, cross-industry changes in portfolio occur if the connections between the industries are sufficiently strong.

The strength of connections between firms in the analyst coverage network also helps us understand the economics of how investors revise their beliefs in response to stock recommendations. We show that the precision of expected return differences between two stocks is increasing in the strength of the connection between these stocks. In addition, if two stocks exhibit a strong connection, then an upward analyst revision to one of these stocks can have a positive spillover effect on the expected excess return on the other stock. To gain further insights on the relation between risk-premia and the structure of the analyst coverage network, we consider the special case in which investors use a one-factor asset pricing model to compute expected returns. We show that the risk-premia and the network can interact in such a way that the sensitivity of returns to changes in the stock recommendations depends
on the stocks factor beta of the stocks. In particular, the sensitivity of expected returns to stock recommendations is higher whenever an analyst revises upwards the value of a stock that has a higher factor exposure than the average factor exposure of all other stocks covered by the analyst.

Finally, our model also provides a novel approach for aggregating analyst recommendations. Traditionally, the consensus stock recommendation has been calculated using the simple average of analyst recommendations (Barber (2001), Jegadeesh (2004)). However, using a simple average completely ignores all the information about the precision of analyst recommendations stored in the second moment. We derive an unbiased, efficient GLS estimator of the consensus analyst recommendation that incorporates both first and second moment information. We show that even if analyst recommendations are unbiased, our estimator asymptotically outperforms the simple average in all cases. The simple average only coincides with our efficient estimator when analyst recommendations are unbiased and analyst recommendations are very imprecise. Thus, our model provides a new approach for estimating consensus recommendations that may be useful for future empirical studies.

Our paper contributes to several strands of literature. First, our paper highlights an intricate link between the literature on Bayesian Portfolio Choice (Black and Litterman (1992), Zhou (2009), Goffman and Manela (2011)) and the literature on the use of graph theory and networks in finance (Anton and Polk (2013), DeGroot (1974), Golub and Jackson (2010), Kelly et al. (2013)). We show how the analyst coverage network impacts information aggregation and portfolio choice in a Bayesian setting. In the optimal Bayesian investment strategy, reallocation across industries depends exclusively on the structure of the Laplacian matrix of the analyst coverage network. In addition, the strength of the connections within the network determines how to adjust the weights in the optimal portfolio in response to changes in analyst recommendations. One of the key contributions of our paper is to show that the structure of the coverage network provides the information necessary for this weighting on the information. Second, our paper develops a portfolio approach that
mitigates the known bias in analyst recommendations. A large literature documents that analyst recommendations may be biased because of career concerns (Hayes (1998), Hong and Kubik (2003)), investment banking relationships (Michaely and Womack (1999), Kadan et al. (2009)), and preferences for stocks with certain quantitative characteristics (Jegadeesh (2004)). Our paper provides a formal method for efficiently extracting information about excess future returns even when analyst recommendations display systematic biases. Third, our model helps explain some of the extant empirical findings related to analyst stock recommendations. Boni and Womack (2006) show that analysts create value only by ranking stocks within industries. Jegadeesh et al. (2004) find that the level of the consensus analyst recommendation contains no marginal predictive power about returns. In other words, the extant literature finds that the value of analyst recommendations comes from their ability to rank stocks relatively rather than absolutely. Our model admits this empirical finding. When analyst recommendations are biased and investors have uninformative priors, only relative valuations matter. Moreover, when all industries belong to disconnected components of the analyst coverage network, the optimal portfolio only reallocates wealth relatively among stocks within industries. Reallocation across industries only occurs when industries are “bridged” by a common analyst. This supports the ideas in Kadan et. al (2012) and Boni and Womack (2006) that firm recommendations only contain information about industry level prospects when analysts use a market benchmark.

The rest of the paper is organized as follows. Section 1 describes the base model and solves the optimal Bayesian portfolio problem. Section 2 discusses the analyst coverage network and the reallocation of wealth in the optimal Bayesian Portfolio. Section 3 discusses the relationship between risk-premia and the structure of the analyst coverage network for a one-factor model of expected returns. Section 4 concludes.
1 A Model Of Analyst Recommendations

A mean-variance investor seeks to choose the portfolio that maximizes his expected utility conditional on information from stock recommendations. There are \( N \) tradable stocks labeled \( i = 1, ..., N \) in the economy with realized next period returns \( r_i \). The investor is assumed to have an asset pricing model to form return expectations, \( E(r_i) \), and uses the model and any data to estimate the covariance matrix of realized returns \( \text{var}(r) = \Sigma \). The realized return for each stock \( i \) is expressed as a function of the expected return and two mean zero random variables \( \theta_i \) and \( \zeta_i \):

\[
    r_i = E(r_i) + \theta_i + \zeta_i
\]

where \( \theta_i \) is the learnable component of return variation and \( \zeta_i \) is the unlearnable component. Our focus is on inferring the learnable component of return variation from analyst stock recommendations.

The variable \( \theta_i \) is intended to capture the variation in fundamentals that analysts can learn through research, and \( \theta_i \) is not public knowledge. The shock \( \zeta_i \) captures risk fundamentals that cannot be learned before the returns are realized. By construction, \( \theta_i \) and \( \zeta_i \) are independent and have zero means. In addition, we assume that only a fraction \( \tau \in [0, 1] \) of the unconditional return variation can be learned, so that:

\[
    \text{var}(\zeta) = (1 - \tau)\Sigma
\]

\[
    \text{var}(\theta) = \tau \Sigma
\]

We define the excess return of stock \( i \) as:

\[
    R_i = r_i - E(r_i) = \theta_i + \zeta_i
\]
We now introduce a total of $A$ stock analysts labeled $a = 1, ..., A$. Analysts provide recommendations about the stocks they cover and we assume their recommendations are a signal of the learnable component of returns $\theta$. Denote by $N_a$ (and by $N_a$) the set (number) of firms covered by analyst $a$ and by $A_a$ (and by $A_a$) the set (number) of analysts that covers firm $i$. For each of the $i \in N_a$ stocks covered by analysts $a$, the analysts reports a noisy signal $y_{i,a}$ about the learnable component of the excess returns of the stock:

$$y_{i,a} = \theta_i + u_a + \varepsilon_{ia} \text{ for all } a = 1, ..., A \text{ and } i \in N_a$$

The unobserved term $u_a$ is an analyst specific bias or measurement error that is common across all stocks that the analyst covers. The noisy component $\varepsilon_{ia}$ is such that $E(\varepsilon_{ia}) = 0$. We allow for a general covariance matrix $\text{var}(\varepsilon) = \Omega$, where $\Omega$ is a positive definite matrix. All the signals can be compactly written in matrix form as

$$y = X\theta + Bu + \varepsilon,$$

where the vectors $y$, $\theta$, $u$ and $\varepsilon$ and the indicator matrices $X$ and $B$ are defined below.

Consider any ordering of the stocks so that $\iota(j,a) \in N_a$ denotes the $j$-th stock covered by analyst $a$ where $j = 1, ..., N_a$ (the ordering is a bijection $\iota(\cdot, a) : \{1, ..., N_a\} \to N_a$). We can represent all the stock recommendations, $y$, in vector notation as the $\sum_{a \in A} N_a$ dimensional column vector where the stock recommendation of all analysts from $a = 1$ to $A$ are stacked-up following the ordering $\iota(\cdot, a)$:

$$y'_a = \left[ y_{\iota(1,a)a}, ..., y_{\iota(N_a,a)a} \right] \text{ and } y' = \left[ y'_1, ..., y'_A \right].$$

The vector $\varepsilon$ with the errors is defined similarly as:

$$\varepsilon'_a = \left[ \varepsilon_{\iota(1,a)a}, ..., \varepsilon_{\iota(N_a,a)a} \right] \text{ and } \varepsilon' = \left[ \varepsilon'_1, ..., \varepsilon'_A \right].$$
and we define the matrix $\Omega = E(\varepsilon \varepsilon')$ with $E(\varepsilon) = 0$. The vectors $\theta$ and $u$ are column vectors with the firm-returns and the bias/measurement error, defined by $\theta' = [\theta_1, ..., \theta_N]$ and $u' = [u_1, ..., u_A]$.

The vectors $\theta$ and $u$ are column vectors with the rm-returns and the bias/measurement error, defined by $\theta' = [\theta_1, ..., \theta_N]$ and $u' = [u_1, ..., u_A]$. The stocks covered by analyst $a$ are represented by the indicator (dummy) matrix $X_a$. The matrix $X_a$ has dimension $N_a \times N$ and, for all $j = 1, ..., N_a$ and $i = 1, ..., N$, $X_a(j, i) = 1$ if $i = \iota(j, a)$ and is otherwise equal to zero. The stock coverage of all analysts are denoted by the indicator matrix $X$. The matrix $X$ has dimension $\left(\sum_{a \in A} N_a\right) \times N$ and is obtained by stacking up all the matrices $X_a$, i.e. $X' = [X'_1, ..., X'_A]$. Note that the rows of $X$ are indexed by $ja$ where $a = 1, ..., A$ and $j = 1, ..., N_a$, and the columns are indexed by $i = 1, ..., N$, and $X(ja, i) = 1$ if $i = \iota(j, a)$ and $X(ja, i) = 0$ otherwise.

Similarly, the matrix $B$ is the analyst indicator matrix with dimension $\left(\sum_{a = 1}^A N_a\right) \times A$. The rows of $B$ are indexed by $ja$ where $a = 1, ..., A$ and $j = 1, ..., N_a$ and the columns are indexed by $b = 1, ..., A$ and $B(ja, b) = 1$ if $a = b$ and $B(ja, b) = 0$ otherwise.

### 1.1 Portfolio Selection

We now solve for the optimal Bayesian portfolio. After observing the stock recommendations signal $y$, investors update their prior beliefs about the return vector $r$ and compute the posterior distribution of returns. This posterior distribution is a normal distribution with mean $E(r|y)$ and covariance matrix $var(r|y)$. Once this posterior distribution is obtained, the investor’s problem is to choose the portfolio weights $\omega = (\omega_i)_{i \in \mathcal{N}}$ that maximizes the following expected utility function conditional on the signal of analyst recommendations:

$$E[U(\omega|y)] = \omega' (E(r|y) - r_f) - \frac{\gamma}{2} \omega' var(r|y) \omega$$

where $r_f$ is the risk-free rate of return. The solution to this problem is well-known and is given by:
ω = \frac{1}{\gamma} \left( \text{var} \,(r|y) \right)^{-1} (E \,(r|y) - r_f) = \frac{1}{\gamma} \left( \text{var} \,(R|y) \right)^{-1} (E \,(r) - r_f + E \,(R|y))

1.1.1 Benchmark Case - Unbiased Stock Recommendations

As a benchmark to compare our results, consider the situation where stock recommendations are not biased. Therefore, stock recommendations are given by

\[ y = X\theta + \varepsilon \text{ where } E(\varepsilon) = 0 \text{ and } \text{var}(\varepsilon) = \Omega. \]

Applying Bayes rule, we obtain the following result which closely resembles Black and Litterman (1992).

**Proposition 1** When investors observe unbiased stock recommendations \( y = X\theta + \varepsilon \) where \( E(\varepsilon) = 0 \) and \( \text{var}(\varepsilon) = \Omega \), the investor obtains the updated expected returns and variance

\[
E \,(R|y) = E \,(\theta|y) = \left( (\tau \Sigma)^{-1} + X'\Omega^{-1}X \right)^{-1} X'\Omega^{-1}y,
\]

\[
\text{var} \,(R|y) = \left( (\tau \Sigma)^{-1} + X'\Omega^{-1}X \right)^{-1} + (1 - \tau) \Sigma,
\]

**Proof:** All proofs are contained in the appendix.

Interestingly, when analyst recommendations are unbiased we have that absolute analyst recommendations carry investment value. This is because the recommendations for the individual stocks are directly comparable across analysts if recommendations are not biased. In other words, for two analysts \( a \) and \( b \), the recommendation \( y_{ia} \) is directly comparable to \( y_{ib} \). The only valuation differences that arise are because of the random error component. However, this is not the case when analyst recommendations are biased. Since each recommendation includes an analyst-specific bias component, the investor is unable to disentangle the differences between \( y_{ia} \) and \( y_{ib} \) that are attributable to noise versus bias. Therefore, only relative valuations (e.g. rankings of firms by a single analyst) carry investment value when
recommendations are biased.

1.1.2 Biased Stock Recommendations

We now turn to the general case where stock recommendations may be biased. We assume that the investor has an uninformative (diffuse) prior over all possible biases $u$ on the stock recommendations. Recall also that the prior belief is that $\theta \sim N(0, \tau \Sigma)$, and $\zeta \sim N(0, (1 - \tau) \Sigma)$ and $R = \theta + \zeta$. Upon observing the stock recommendation vector $y$, the investor uses Bayes rule to obtain the posterior distribution for excess returns by integrating over all possible biases $u$ on the stock recommendations. The posterior expected excess return and variance and then given by the following result:

**Proposition 2** Suppose that the stock recommendations are given by $y = X\theta + Bu + \varepsilon$, where $E(\varepsilon) = 0$ and $\text{var}(\varepsilon) = \Omega$. The expectation and variance of returns conditional on the stock recommendations, $E(R|y)$ and $\text{var}(R|y)$, are given by

$$E(R|y) = E(\theta|y) = \left( (\tau\Sigma)^{-1} + X'QX \right)^{-1} X'Qy,$$

and

$$\text{var}(R|y) = \left( (\tau\Sigma)^{-1} + X'QX \right)^{-1} + (1 - \tau) \Sigma,$$

$$\text{var}(\theta|y) = (\tau\Sigma)^{-1} + X'QX^{-1},$$

where $Q = \Omega^{-1} - \Omega^{-1}B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1}$.

Note that when analyst recommendations are biased, the precision of the relative stock recommendations, $\text{var}(\theta|y)$, is lower than in the case of unbiased recommendations. The total loss in precision is given by $X'\Omega^{-1}B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1}X$. The optimal Bayesian portfolio is then calculated directly using Proposition 2 and is summarized by the following proposition.
The Sharpe ratio of the investment strategy is equal to \( Sh = \sqrt{E(R|y)^\prime \text{var}(R|y)^{-1} E(R|y)} \) and, as should be expected, is the maximum Sharpe ratio among all feasible investment strategies.

**Proposition 3** Investors adjust their portfolio weights by \( \Delta \omega = \frac{1}{\gamma} (\text{var}(R|y))^{-1} E(R|y) \) conditional on receiving the stock recommendation where

\[
\Delta \omega = \frac{1}{\gamma} \left( \left( (\tau \Sigma)^{-1} + X'QX \right)^{-1} + (1 - \tau) \Sigma \right)^{-1} \left( (\tau \Sigma)^{-1} + X'QX \right)^{-1} X'Qy
\]

where \( \text{var}(\varepsilon) = \Omega \) and \( Q = \Omega^{-1} - \Omega^{-1} B (B'\Omega^{-1} B)^{-1} B'\Omega^{-1} \) and the portfolio weights are

\[
\omega = \Delta \omega + \frac{1}{\gamma} (\text{var}(R|y))^{-1} (E(r) - r_f).
\]

It is worth noting that we can express the optimal weights by: \( \omega = \Delta \omega + \omega^e \), where \( \omega^e = \frac{1}{\gamma} (\text{var}(R|y))^{-1} (E(r) - r_f) \) are the portfolio weights when \( E(R|y) = 0 \) (or how much investors would invest when receiving a zero posterior signal). We are mainly interested in the changes in portfolio choice, \( \Delta \omega \), in response to stock recommendations. In the next section we show how \( \Delta \omega \) is intricately related to the structure of the analyst coverage network. For now though, we show how \( \Delta \omega \) depends critically on the information environment. First, consider altering the precision of analyst information. In the case where the analyst information is very imprecise: \( \Omega = \sigma^2 I \) and \( \sigma \to \infty \), we have that \( E(R|y) \to 0 \), \( \text{var}(R|y) \to \Sigma \), and the portfolio choice is unchanged by the stock recommendations so \( \omega = \omega^e \). On the other hand, in the case where analyst information is extremely precise: \( \Omega = \sigma^2 I \) and \( \sigma^{-1} \to \infty \), we have that all the learnable component should be learned. Second, consider making all the information learnable by setting \( \tau = 1 \). In this case the portfolio choice takes on a particularly simple form \( \Delta \omega = \frac{1}{\gamma} X'Q \). The investment rule that arises implies that investors apply more weight when they have multiple favorable demeaned information, and large stocks are much more affected than small stocks with low following by analysts.
The posterior moments under unbiased and biased recommendations differ only by the interaction of the matrix $X$ with the matrix $Q$. It turns out that the matrix $X'QX$ plays a central role in our analysis as it embodies the structure of the data. This matrix is intricately linked to the structure of the analyst coverage network and how wealth is reallocated across the portfolio. We will turn to the study of its properties in the next section, but we collect in the following result describing its properties.

**Lemma 1** The matrix $X'QX$ where $Q = \Omega^{-1} - \Omega^{-1}B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1}$ has the following properties:

(i) $X'QX$ is a symmetric positive semi-definite matrix.

(ii) The rank of $X'QX$ is equal to $N - K$ where $K$ is the number of connected components of the graph $G$.

(iii) The rows and column sums of $X'QX$ and $Q$ are zero.

### 1.2 Unbiased Consensus Stock Recommendations

We now consider the problem of estimating the unbiased consensus stock recommendations. In the model, stock recommendations reveal information about excess returns. Stock recommendations and excess returns are related by the regression of the stock recommendations on the dummy variables $X$ and $B$, $y = X\theta + Bu + \varepsilon$.

We first estimate the unbiased consensus stock recommendations when analyst recommendations are not biased. Therefore, stock recommendations and excess returns are related by the regression $y = X\theta + \varepsilon$. The unbiased, efficient estimator of the consensus stock recommendations is simply given by the GLS estimator of $\theta$:

$$\hat{R} = (X'\Omega^{-1}X)^{-1}X\Omega^{-1}y$$

Note that this estimator includes information contained in both the first and the second
moment of the distribution. This is in contrast to measure of the consensus analyst recommendation that is normally used in empirical work. Most empirical studies (Barber (2001), Jegadeesh (2004), among others) adopt the simple average of analyst recommendations as their estimator. While the simple average is unbiased in this setting, it is asymptotically inefficient because it ignores all the precision information stored in the second moment. Moreover, only in the case where $\Omega = \sigma^2 I$ does our unbiased estimator coincide with the average stock recommendation:

$$\hat{R}_i = \frac{1}{|A_i|} \sum_{a \in A_i} y_{ia}. $$

where $A_i$ is the set of all analysts covering stock $i$.

Now consider the problem of estimating the unbiased consensus stock recommendations when analyst recommendations are biased. The simple average will no longer be an unbiased estimator of excess returns if the bias vector $u$ has non-zero mean. Thus, we need to directly estimate the $\theta$ parameter via the regression $y = X\theta + Bu + \epsilon$ if we want to obtain an unbiased estimator of excess returns. An econometric issue arises in this regression because not all $N$ coefficients $\theta$ and $A$ coefficients $u$ in the regression above can be identified because the rank of the regressor $[X|B]$ is less than $N + A$. The columns of the matrix $[X|B]$ are linearly dependent because the sum of all columns of $X$ and the sum of all columns of $B$ are the vectors of ones. This is not the only source of linear dependency. In particular, the rank of $[X|B]$ is $N + A - K$ ($K \geq 1$), where $K$ reflects the number of connected components of the analyst coverage network. This concept is discussed in more detail later in the paper.

In order to identify the parameters in the regression, we can drop a firm inside each connected component and estimate all returns relative to the dropped firm. Alternatively, we can add a restriction on the coefficients $\sum_{i \in \mathcal{N}(k)} \theta_i = 0$ so that they add up to zero. We prefer the latter approach because it preserves the symmetry of the problem.

We estimate the consensus stock recommendation by GLS using the following constrained
\[
y = X\theta + Bu + \varepsilon
\]

\[
\sum_{i \in \mathcal{N}^{(k)}} \theta_i = 0 \text{ for all } k = 1, \ldots, K
\]

**Proposition 4** The unbiased efficient estimator of the consensus stock recommendations can be obtained from the analyst stock recommendations using the GLS estimator

\[
\hat{R}_{\text{GLS}} = (X'QX)^+ X'y, \text{ with variance } \text{var}(\hat{R}_{\text{GLS}}) = (X'QX)^+,
\]

and precision \(\text{var}(\hat{R}_{\text{GLS}})^+ = X'QX\), where \((X'QX)^+\) is the Moore-Penrose generalized inverse of \(X'QX\), \(Q = \Omega^{-1} - \Omega^{-1}B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1}\), and \(\text{var}(\varepsilon) = \Omega\).

## 2 The Analyst Coverage Network

So far we have calculated the optimal Bayesian portfolio and the estimator for the unbiased consensus stock recommendations when analyst recommendations are biased. We now seek to better understand the economics of our results, with particular emphasis on understanding the matrix \(X'QX\) that appears in both the optimal Bayesian portfolio calculation and the unbiased estimator formula. In this section we show how \(X'QX\) is intricately linked to how wealth is reallocated among the optimal portfolio upon observing the signal \(y\).

### 2.1 Definitions

The *analyst coverage network* is defined as the graph \(G\), where the vertices of the graph are the firms \(N\) and the edges are all the pairs of distinct firms that are both covered by at least one common analyst. Formally, let \(\mathcal{A}_{ij} = \{a \in \mathcal{A} : i \in N_a \text{ and } j \in N_a\}\) be the set of all analysts covering both firms \(i\) and \(j\). The graph \(G\) consists of the set of vertices \(V(G) = N\) and the set of edges \(E(G) = \{\{i, j\} \in N \times N : i \neq j \text{ and } \mathcal{A}_{ij} \neq \emptyset\}\). We use the notation
$i \sim j$ to indicate that two firms $i$ and $j$ are joined by an edge, i.e., there is an analyst in common following both firms.

Two firms $i$ and $j$ are defined as connected if and only if there is a path connecting them. That is, there is a set of firms $i = i_0, i_1, i_2, \ldots, i_m = j$ such that $i_{k-1} \sim i_k$ for $k = 1, \ldots, m$.

Connection is an equivalence relation. We will refer to the components of the graph as the maximal equivalence classes of the connection relation. Furthermore, a graph is defined as connected if any two firms can be joined by a path and otherwise is disconnected. A maximum connected subgraph of the graph is defined as a component.

To highlight the above ideas, we present the following result that establishes an important relationship between the regression in Section 1.3 and the analyst coverage network.

**Proposition 5** Let the rank of the matrix $[X|B]$ be $N + A - K$. Then the following holds:

(i) $K = N + A - \text{rank}([X|B]) \geq 1$ and $K$ is the number of connected components of the graph $G$;

(ii) The set of firms $\mathcal{N}$ and the set of analysts $\mathcal{A}$ can be partitioned into $K$ connected (disjoint) components $\mathcal{N}^{(k)}$ and $\mathcal{A}^{(k)}$ such that $\mathcal{N} = \bigcup_{k=1}^{K} \mathcal{N}^{(k)}$, $\mathcal{A} = \bigcup_{k=1}^{K} \mathcal{A}^{(k)}$ and such that any two stocks $i$ and $j$ in $\mathcal{N}^{(k)}$ are connected by analysts belonging to $\mathcal{A}^{(k)}$.

By imposing a restriction on the error terms, we can obtain $K$ completely unrelated regressions corresponding to the stocks in each of the $K$ connected components, $\mathcal{N}^{(k)}$, and analysts, $\mathcal{A}^{(k)}$, which can be estimated separately.

**Corollary 1** Let $K = N + A - \text{rank}([X|B])$. If the errors $\varepsilon_{ia}$ are clustered by firms and/or by analysts, then there are $K$ completely unrelated regressions corresponding to the regression of the analyst stock recommendations for the stocks in each of the $K$ connected components of the graph $G$ on the dummy variables for the stocks in the connected component and the analysts that covers those stocks.

**Proof:** The errors are clustered by firms and by analysts so by our assumption on $\Omega$, the $\text{cov}(\varepsilon_{ia}, \varepsilon_{jb}) = 0$ for any $i \in \mathcal{N}^{(k)}$ and $a \in \mathcal{A}^{(k)}$ and $j \in \mathcal{N}^{(k')}$ and $b \in \mathcal{A}^{(k')}$. 

15
2.2 Wealth Reallocation Across Components

We now show how wealth is reallocated among stocks in the optimal portfolio. The following proposition summarizes one of the main results of the paper:

**Proposition 6** There is no reallocation of wealth among disconnected components.

Proposition 6 implies that the reallocation of wealth occurs only within connected components of the analyst coverage network. Even though returns across stocks in disconnected components may be correlated, there is no overall reallocation of wealth among any disconnected components. This result directly relates to the concept of relative valuation. Since stock analyst recommendations are biased, absolute valuation is worthless and investors can only make inferences about the relative value of stocks that are connected. At the component level this implies that the informational content of stock recommendations only carries value within the components. Even if the stocks is one component display systematically higher ratings than a separate, disconnected, component, investors cannot infer whether this occurs because the stocks have higher expected returns or if the analysts in the original component are just more optimistic. Only when there is a connection between these groups can inferences be made and wealth reallocated.

2.3 Relative Wealth Reallocation

We now show how wealth is reallocated relatively within components by studying the strength of the connections in the analyst coverage network. To simplify the analysis going forward, we will maintain throughout this section the assumption that the errors are heteroscedastic and are pairwise uncorrelated. Let $\text{var}(\varepsilon_{ia}) = \sigma_{ia}^2$ and let the precision of the signal be denoted by $\tau_{ia} = \sigma_{ia}^{-2}$. Therefore, $\Omega = \text{diag}(\sigma_{ia}^2)$. 
2.3.1 The Weighted Analyst Coverage Network

The weighted analyst coverage network describes the strength of the connections throughout the network. For any two distinct connected firms $i \sim j$ define the weighted graph $G$ by the following adjacency matrix

$$A(G)_{ij} = \sum_{a \in A_{ij}} \left( \frac{\tau_{ia} \tau_{ja}}{\sum_{k \in N_a} \tau_{ka}} \right), \text{ where } \tau_{ia} = var(\varepsilon_{ia})^{-1},$$

and set $A(G)_{ij} = 0$ for any two firms $i$ and $j$ that are not connected (also $A(G)_{ii} = 0$). The matrix $A(G)$ is the adjacency matrix of the graph $G$. The entries in the adjacency matrix captures the relative strength of the connections between firms.

Furthermore, define the degree of a vertex in the graph $G$ as:

$$d(i) = \sum_{j \in N} A(G)_{ij}$$

and the degree matrix of graph $G$ is defined as the diagonal matrix with the degrees in the diagonal,

$$D(G) = diag(d(1), ..., d(N)).$$

The degree increases when there are more analysts covering a stock and when the analyst covering the stock cover more stocks. The contribution to the degree of an analyst that covers more stocks is higher than an analyst that covers less stocks.

The key matrix representing the properties of the graph is the Laplacian matrix $L(G)$, defined by:

$$L(G) = D(G) - A(G)$$

The Laplacian turns out to be equal to the matrix $X'QX$, the matrix that describes the differences between the benchmark portfolio and the optimal portfolio with biased recommendations. The rank of $L(G)$ is equal to $N - K$ where $K$ is the number of connected
components of the graph $G$. Formally:

**Proposition 7** The Laplacian matrix is equal to $L(G) = X'QX$ where $\text{var}(\varepsilon) = \Omega = \text{diag}(\sigma_{\varepsilon}^2)$ and $Q = \Omega^{-1} - \Omega^{-1}B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1}$.

The Laplacian matrix plays a key role in our analysis because, as we shall show next section, it is the precision matrix of the unbiased stock recommendations. The matrix captures the estimation risk in the stock recommendation signals.\(^1\)

### 2.3.2 The Economics of Relative Valuations

Our goal is to understand the economics of how wealth is reallocated among components and the economics of the expressions we have derived. We start with the simple case in which $\Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_N^2)$ and the stocks are not pairwise correlated. The following lemma describes some of the important properties of the posterior variance (precision matrix) that will be used later on:

**Lemma 2** Let $\Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_N^2)$. Then the variance matrix $\text{var}(R|y) = V + (1 - \tau)\Sigma$ where $V = ((\tau \Sigma)^{-1} + D - A)^{-1}$ and $A$ is the adjacency matrix and $D$ is the degree matrix of the graph, satisfy the following properties:

(i) The matrix $V$ is non-negative and the diagonal entry is greater than any off-diagonal entry: $V_{ii} \geq V_{ij} \geq 0$ for all $i$ and $j$;

(ii) The matrix $V = ((\tau \Sigma)^{-1} + D - A)^{-1} = \Theta^{-1/2} \left( \sum_{k=0}^{\infty} (\Theta^{-1/2}A\Theta^{-1/2})^k \right) \Theta^{-1/2}$, where $\Theta = (\tau \Sigma)^{-1} + D$ is a diagonal matrix.

Let the normalized adjacency matrix of the graph be defined by:

$$\tilde{A} = \Theta^{-1/2}A\Theta^{-1/2}$$

\(^1\)The Laplacian matrix plays an important role in graph theory. The eigenvalues of the Laplacian matrix are the graph spectrum which is one of the main invariants in graph theory describing the main properties and the structure of a graph. The eigenvalues of the Laplacian play a much more fundamental role in describing the properties of a graph than the eigenvalues of the more familiar adjacency matrix.
The number of paths of length $k$ connecting two vertices $i$ and $j$ is given by $\tilde{A}_{ij}^k$. The sum $\sum_{k=0}^{\infty} \tilde{A}_{ij}^k$ is the total number of paths connecting $i$ and $j$. We can interpret $\sum_{k=0}^{\infty} \tilde{A}_{ij}^k$ as the strength of the connection between $i$ and $j$. In other words, the stronger (weaker) is the connection between two firms $i$ and $j$ the larger (smaller) is the value of

$$V_{ij} = \Theta_{ii}^{-1/2} \left( \sum_{k=0}^{\infty} \tilde{A}_{ij}^k \right) \Theta_{jj}^{-1/2}$$

We now show that the precision of the relative value between stocks $i$ and $j$, $\text{var}(R_i - R_j|y)^{-1}$, is greater when the stocks are more closely connected. Firms which share more common analysts are thus easier to make relative judgments about and easier to assign relative weights in the portfolio to one another.

**Corollary 2** The precision of the relative value between stocks $i$ and $j$, that is $\text{var}(R_i - R_j|y)^{-1}$, is greater (smaller) when the connection between firms $i$ and $j$ is stronger (weaker).

**Proof:** The variance of the relative value between stocks $i$ and $j$ is given by

$$\text{var}(R|y) (\delta_i - \delta_j) = \text{var}(R_i - R_j|y) = \text{var}(R_i|y) + \text{var}(R_j|y) - 2\text{cov}(R_i, R_j|y) = V_{ii} + V_{jj} - 2V_{ij} + (1 - \tau) (\Sigma_{ii} + \Sigma_{jj} - 2\Sigma_{ij})$$

where $V = \text{var}(\theta|y)$. Thus the precision $\text{var}(R_i - R_j|y)^{-1}$ is increasing in $V_{ij}$ which we argued above increases with the strength of the connection of two firms. QED.

The next proposition summarizes our other main result of how wealth is reallocated within components. The following proposition shows that stocks are reallocated more weight (via their posterior expected return increasing) after receiving favorable analyst recommendations. Moreover, if the connection between two stocks $i$ and $j$ is sufficiently strong (weak), a favorable analyst recommendation for stock $i$ may result in a upward (downward) revision to the posterior expected excess return of stock $j$. This implies that reallocation within
components depends on not only the value of relative stock recommendations, but also the strength of connections between stocks in the network.

**Proposition 8** The sensitivity of returns to stock recommendations satisfy

\[
\frac{\partial E(R|y)_j}{\partial y_{ia}} = \tau_{ia} \left( V_{ji} - \sum_{k \in N_a} V_{jk} \bar{\tau}_{ka} \right), \text{ where } \bar{\tau}_{ia} = \frac{\tau_{ia}}{\sum_{k \in N_a} \tau_{ka}}
\]

and \( N_a \) are the stocks covered by analyst \( a \) and \( \tau_{ia} = \text{var}(\varepsilon_{ia})^{-1} \). The following comparative statics results hold:

(i) The stock \( i \) return increase when any analyst following stock \( i \) revise it upwards. That is

\[
\frac{\partial E(R|y)_i}{\partial y_{ia}} \geq 0.
\]

(ii) When the strength of the connection between stock \( j \) and stock \( i \) is stronger (weaker) than the average strength of the connection among stock \( j \) and the other stocks covered by analyst \( a \), then stock \( j \) return increases when an analyst revise stock \( i \) upwards. Formally,

\[
\frac{\partial E(R|y)_j}{\partial y_{ia}} \gtrless 0 \text{ if and only if } V_{ji} \gtrless \sum_{k \in N_a} V_{jk} \bar{\tau}_{ka}.
\]

### 3 Risk-Premia and the Analyst Coverage Network

We finally turn toward understanding the relation between risk-premia and the structure of the analyst coverage network. Consider the following one-factor asset pricing model

\[
R_i = \beta_i f + \varepsilon_i, \text{ where } E(\varepsilon) = 0 \text{ and } \text{var}(\varepsilon) = \text{diag}(\sigma_1^2, \ldots, \sigma_N^2) = \Sigma_\varepsilon, \text{ and } \text{var}(f) = 1.
\]

The unconditional variance of returns are given by:

\[
\text{var}(R) = \Sigma_\varepsilon + \beta \beta'.
\]

We now show that the sensitivity of returns to stock recommendations can be decomposed in two parts. The first part, which we analyzed in the previous section, plus a second part
related how the covariance of returns and interacts with the network structure.

**Proposition 9** Let \( \Sigma = \Sigma_\varepsilon + \beta \beta' \) be the unconditional return variance, where \( \Sigma_\varepsilon \) are the idiosyncratic stock variances and \( \beta \) is the beta of the stocks. Then:

(i) The conditional return variance is

\[
\text{var}(R|y) = V + \kappa b b' + (1 - \tau) (\Sigma_\varepsilon + \beta \beta').
\]

where the values of \( V, b \) and the constant \( \kappa > 0 \) are given by

\[
V = \left( \tau^{-1} \Sigma_\varepsilon^{-1} + X'QX \right)^{-1},
\]

\[
b = \left( \tau^{-1} \Sigma_\varepsilon^{-1} + X'QX \right)^{-1} \Sigma_\varepsilon^{-1} \beta, \text{ and}
\]

\[
\kappa = \left( \tau (1 + \beta' \Sigma_\varepsilon^{-1} \beta) - \beta' \Sigma_\varepsilon^{-1} (\tau^{-1} \Sigma_\varepsilon^{-1} + X'QX)^{-1} \Sigma_\varepsilon^{-1} \beta \right)^{-1} > 0.
\]

(ii) The sensitivity of returns to changes in the stock recommendations by analyst \( a \) satisfy

\[
\frac{\partial E(R|y)_j}{\partial y_{ia}} = \tau_{ia} \left( V_{ji} - \sum_{k \in N_a} V_{jk} \bar{\tau}_{ka} \right) + \kappa \tau_{ia} b_j \left( b_i - \sum_{k \in N_a} b_k \bar{\tau}_{ka} \right),
\]

where \( N_a \) are the stocks covered by analyst \( a \), \( \bar{\tau}_{ia} = \frac{\tau_{ia}}{\sum_{k \in N_a} \tau_{ka}} \) and \( \tau_{ia} = \text{var}(\varepsilon_{ia})^{-1} \).

Note that the second term in the expression for the sensitivity depends on \( b \) and the interaction with the network structure, embedded in \( (\tau^{-1} \Sigma_\varepsilon^{-1} + X'QX) \) and the \( \beta \)'s. Thus, the revision to the posterior expected excess return for stock \( j \) additionally depends on the risk-premia in the model. The following corollary describes the conditions under which the posterior expected excess return coincides with the results from Proposition 8:

**Corollary 3** Let the unconditional returns variance be \( \Sigma = \Sigma_\varepsilon + \beta \beta' \), where \( \Sigma_\varepsilon \) are the idiosyncratic stock variances and all stocks have the same \( \beta \). Then

\[
\frac{\partial E(R|y)_j}{\partial y_{ia}} = \tau_{ia} \left( V_{ji} - \sum_{k \in N_a} V_{jk} \bar{\tau}_{ka} \right),
\]

21
where $N_a$ are the stocks covered by analyst $a$, $\tau_{ia} = \text{var}(\varepsilon_{ia})^{-1}$ and $\overline{\tau}_{ia} = \frac{\tau_{ia}}{\sum_{k \in N_a} \tau_{ka}}$.

The following corollary shows more explicitly that the risk-premia and the network can interact in a way that the sensitivity of the returns to changes in the stock recommendations depends on the betas of the stocks. In particular, the sensitivity of the stock recommendation is going to be higher whenever an analyst revises upwards the value of a stock that has higher beta than the average of the other stocks covered by the analyst.

**Corollary 4** Let the unconditional return variance be $\Sigma = \Sigma_\varepsilon + \beta \beta'$, where $\Sigma_\varepsilon = \sigma_\varepsilon^2 I$ are the idiosyncratic risks and let $\beta = c + dv_2$ where $v_2$ is the Fiedler eigenvector, corresponding to the second smallest eigenvector of the Laplacian matrix $X'QX$, and $c$ and $d$ are two constants. Thus the sensitivity of the stock recommendation is going to be higher whenever analyst $a$ revise upwards the value of a stock that has higher beta than the average of the other stocks covered by the analyst. Formally,

$$
\frac{\partial E(R|y)_j}{\partial y_{ia}} = \tau_{ia} \left( V_{ji} - \sum_{k \in N_a} V_{jk} \overline{\tau}_{ka} \right) + \kappa \tau_{ia} b_j \sum_{k \in N_a} (\beta_i - \beta_k) \overline{\tau}_{ka}
$$

and $\kappa_j > 0$ is given by $\kappa_j = \frac{\kappa}{|N_a|} b_j$.

4 Conclusion

Investors face a difficult challenge when incorporating biased sell-side analyst recommendations into their asset allocation decisions. In this paper, we proposed a formal method for alleviating this challenge in a Bayesian framework. We first solved the optimal Bayesian portfolio problem in the presence of biased analyst recommendations. The optimal portfolio could be cleanly expressed as the sum of two components: how much investors would invest when receiving a zero posterior signal and how much they reallocate among stock in response to stock recommendations. We found that the reallocation of wealth across stocks
was intricately connected to the analyst coverage network - 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. Upon observing analyst recommendations, no wealth is reallocated among disconnected components of the analyst coverage network. Instead, wealth is reallocated relatively within connected components according to the strength of the connections between stocks and the value of the relative recommendations. Finally, we derived a novel estimator for the consensus stock recommendation that incorporates both information about the level and the precision of analyst recommendations.

Our paper should be of interest to both academics and practitioners alike. As stated earlier, over 99% of all U.S. stocks belong to one giant connected component of the analyst coverage network. Our paper highlights the connectedness of the U.S. equity market and takes a first step towards efficiently using the information stored in the analyst coverage network. We encourage future research in this area to further understand the implications of this highly connected informational network.
References


Appendix A: Proofs

**Proof of Proposition 1:** Denote by \( p(\cdot) \) the density function and the conditional density function. Applying Bayes rule,

\[
p(\theta|y) = p(y)^{-1}p(\theta, y) = p(y)^{-1}p(\theta, y) = p(y)^{-1}p(y|\theta)p(\theta).
\]

Using that \( y = X\theta + \varepsilon \) where \( E(\varepsilon) = 0 \) and \( \text{var}(\varepsilon) = \Omega \) and the prior of \( \theta \) is \( N(0, \tau \Sigma) \),

\[
p(\theta|y) = p(y)^{-1}G(X\theta, \Omega, y)G(0, \tau \Sigma, \theta)
\]

Let the \( d \)-dimensional Gaussian density function, where \( \mu \in \mathbb{R}^d \) and \( A \in \mathbb{R}^{dxd} \), be

\[
G(\mu, A, x) = (2\pi)^{-d/2} \det(A)^{-1/2} \exp(-\frac{1}{2} (x - \mu)' A^{-1} (x - \mu))
\]

It is well-known that

\[
G(a, A, Qx)G(b, B, x) = G(c, C, x)G(a + QBQ', Qb)
\]

where

\[
C = (Q'A^{-1}Q + B^{-1})^{-1} \text{ and } c = C(Q'A^{-1}a + B^{-1}b)
\]

Obviously, \( G(X\theta, \Omega, y) = G(y, \Omega, X\theta) \). Applying the result above

\[
G(y, \Omega, X\theta)G(0, \tau \Sigma, \theta) = G(c, C, \theta)G(y, \Omega + X\tau \Sigma X', 0)
\]

where

\[
C = (X'\Omega^{-1}X + (\tau \Sigma)^{-1})^{-1}
\]

\[
c = (X'\Omega^{-1}X + (\tau \Sigma)^{-1})^{-1} X'\Omega^{-1}y
\]
Replacing the expression on conditional density,

\[ p(\theta|y) = p(y)^{-1}G(c, C, \theta)G(y, \Omega + X\tau\Sigma X', 0). \]

But since \( \int p(\theta|y)d\theta = 1 \), then \( p(y)^{-1} = G(y, \Omega + X\tau\Sigma X', 0) \), and \( p(\theta|y) = G(c, C, \theta) \). This implies that,

\[ E(\theta|y) = c = (X'\Omega^{-1}X + (\tau\Sigma)^{-1})^{-1}X'\Omega^{-1}y \]
\[ \text{var}(\theta|y) = C = (X'\Omega^{-1}X + (\tau\Sigma)^{-1})^{-1} \]

and then obviously,

\[ E(R|y) = (X'\Omega^{-1}X + (\tau\Sigma)^{-1})^{-1}X'\Omega^{-1}y \]
\[ \text{var}(R|y) = (X'\Omega^{-1}X + (\tau\Sigma)^{-1})^{-1} + (1 - \tau)\Sigma \]

which completes the proof.

QED.

Proof of Proposition 2: Our aim is to obtain the density function of \( p(R|y) \). Applying Bayes rule

\[ p(\theta|y) = p(y)^{-1} \int p(\theta, u, y)du \]
\[ p(\theta|y) = p(y)^{-1} \int p(y|\theta, u)p(\theta, u)du \]

Since \( u \) and \( \theta \) are independent, \( p(\theta, u) = p(\theta)p(u) \), and \( u \) has an uninformative prior, so \( p(u) \) are all equally likely,

\[ p(\theta|y) \propto \int p(y|\theta, u)p(\theta)du \]
Replacing the expressions for the densities $p(y|\theta, u)$ and $p(\theta)$,

$$p(\theta|y) \propto \int \left( \exp \left( -\frac{1}{2} \left( (y - (X\theta + Bu))\Omega^{-1}(y - (X\theta + Bu)) + \theta'(\tau\Sigma)^{-1}\theta \right) \right) \right) du$$

To integrate the expression over $u$ consider the expression inside the exponential function,

$$((y - (X\theta + Bu))\Omega^{-1}(y - (X\theta + Bu)) + \theta'(\tau\Sigma)^{-1}\theta) = (y - X\theta)'\Omega^{-1}(y - X\theta) + (\theta)'(\tau\Sigma)^{-1}\theta + 2(X\theta - y)'\Omega^{-1}(Bu) + (Bu)'\Omega^{-1}(Bu)$$

The only term that is a function of $u$ is $2(X\theta - y)'\Omega^{-1}(Bu) + u'B'^{-1}u$, which can be rewritten as

$$\left( u - (B'^{-1}B)^{-1}B'^{-1}(y - X\theta) \right)'B'^{-1}B \left( u - (B'^{-1}B)^{-1}B'^{-1}(y - X\theta) \right) - (y - X\theta)'\Omega^{-1}B(B'^{-1}B)^{-1}B'^{-1}(y - X\theta)$$

But we know that the integral

$$\int \left( \exp \left( -\frac{1}{2} \left( u - (B'^{-1}B)^{-1}B'^{-1}(y - X\theta) \right)'B'^{-1}B \left( u - (B'^{-1}B)^{-1}B'^{-1}(y - X\theta) \right) \right) \right) du$$

is a constant and does not depend on $\theta$. Therefore the integral is simply

$$p(\theta|y) \propto \exp \left( -\frac{1}{2} \left( (y - X\theta)'\Omega^{-1}(y - X\theta) + \theta'(\tau\Sigma)^{-1}\theta - (y - X\theta)'\Omega^{-1}B(B'^{-1}B)^{-1}B'^{-1}(y - X\theta) \right) \right)$$

In order to further simplify this expression let

$$Q = \left( \Omega^{-1} - \Omega^{-1}B(B'^{-1}B)^{-1}B'^{-1} \right)$$
The following steps show that

\[
\left(y - X\theta\right)\Omega^{-1} \left(y - X\theta\right) + \theta' \left((\tau \Sigma)^{-1} \theta - (y - X\theta)' \Omega^{-1} B (B'\Omega^{-1} B)^{-1} B'\Omega^{-1} (y - X\theta)\right) \\
= \theta' \left((\tau \Sigma)^{-1} + X'QX\right) \theta - 2\theta' X'Qy + y'Qy
\]

Consider the following equalities,

\[
(y' - \theta'X')\Omega^{-1} (y - X\theta) + \theta' (\tau \Sigma)^{-1} \theta - (y' - \theta'X')\Omega^{-1} B (B'\Omega^{-1} B)^{-1} B'\Omega^{-1} (y - X\theta) \\
= \theta' X'\Omega^{-1} X\theta + \theta' (\tau \Sigma)^{-1} \theta - \theta' X'\Omega^{-1} B (B'\Omega^{-1} B)^{-1} B'\Omega^{-1} X\theta \\
- 2\theta' X'\Omega^{-1} y + 2\theta' X'\Omega^{-1} B (B'\Omega^{-1} B)^{-1} B'\Omega^{-1} y + y' \left(\Omega^{-1} - \Omega^{-1} B (B'\Omega^{-1} B)^{-1} B'\Omega^{-1}\right) y \\
= \theta' \left((\tau \Sigma)^{-1} + X' \left(\Omega^{-1} - \Omega^{-1} B (B'\Omega^{-1} B)^{-1} B'\Omega^{-1}\right) X\right) \theta \\
- 2\theta' X' \left(\Omega^{-1} - \Omega^{-1} B (B'\Omega^{-1} B)^{-1} B'\Omega^{-1}\right) y + y' \left(\Omega^{-1} - \Omega^{-1} B (B'\Omega^{-1} B)^{-1} B'\Omega^{-1}\right) y \\
= \theta' \left((\tau \Sigma)^{-1} + X'QX\right) \theta - 2 (\theta)' X'Qy + y'Qy
\]

Now we show that

\[
\left(\theta - ((\tau \Sigma)^{-1} + X'QX)^{-1} X'Qy\right)' \left((\tau \Sigma)^{-1} + X'QX\right) \left(\theta - ((\tau \Sigma)^{-1} + X'QX)^{-1} X'Qy\right) \\
= \theta' \left((\tau \Sigma)^{-1} + X'QX\right) \theta - 2\theta' X'Qy + y'QX \left((\tau \Sigma)^{-1} + X'QX\right)^{-1} X'Qy
\]
Indeed,

\[
(\theta - ((\sigma^{-1} + X'QX)^{-1} X'Qy)') ((\sigma^{-1} + X'QX) (\theta - ((\sigma^{-1} + X'QX)^{-1} X'Qy) = \theta' ((\sigma^{-1} + X'QX) \theta - 2\theta' ((\sigma^{-1} + X'QX) ((\sigma^{-1} + X'QX)^{-1} X'Qy + + y'QX (\sigma^{-1} + X'QX)^{-1} ((\sigma^{-1} + X'QX)((\sigma^{-1} + X'QX)^{-1} X'Qy) = \theta' ((\sigma^{-1} + X'QX) \theta - 2(\theta')' X'Qy + y'QX ((\sigma^{-1} + X'QX)^{-1} X'Qy

The two expressions above show that

\[
p(\theta|y) \propto \exp \left( -\frac{1}{2} ((\theta - \mu)' V^{-1} (\theta - \mu) \right),
\]

where

\[
\mu = ((\sigma^{-1} + X'QX)^{-1} X'Qy
\]

and

\[
V = ((\sigma^{-1} + X'QX)^{-1}
\]

which completes the proof.

QED

Proof of Proposition 5:

Item (ii): Let \( N^{(k)} \) represent the set of firms in each of the \( K \) connected components. The set of firms \( N \) is partitioned into the \( K \) disjoint components \( N^{(k)} \), and \( N = \bigcup_{k=1}^{K} N^{(k)} \).

Now define \( A^{(k)} = \bigcup_{i \in N^{(k)}} A_i \), where \( A_i \) is the set of analysts that cover firm \( i \).

It is clear that the \( A^{(k)} \) also forms a partition of the set of analysts \( A \). Indeed, \( A = \bigcup_{k=1}^{K} A^{(k)} \) and each \( A^{(k)} \) are disjoint. Suppose there was an \( a \in A^{(k)} \cap A^{(k')} \) for two distinct components \( k \) and \( k' \). Then there would exist a firm \( i \in N^{(k)} \) followed by analyst \( a \) and a
firm $i \in \mathcal{N}^{(k)}$ also followed by analyst $a$, which would be a contradiction with $i$ and $j$ not being connected. Moreover, all the connections among firms in $\mathcal{N}^{(k)}$ are exclusively with analysts in $\mathcal{A}^{(k)}$. Indeed, say that $i$ and $j$ in $\mathcal{N}^{(k)}$ are connected by an analyst $a$ belonging to $\mathcal{A}^{(k')}$. Then there exists a firm $q \in \mathcal{N}^{(k')} \cup \mathcal{A}^{(k')}$ such that $a \in \mathcal{A}_q$. But that implies that $i$ and $j$ are also connected to $q$ which is a contradiction. This completes the proof of item (ii).

Item (i): The sum of all columns of $\sum_{j=1}^{\mathcal{N}} X_j = 1$, the vectors of ones and also $\sum_{a=1}^{\mathcal{A}} B_a = 1$, thus $K \geq 1$.

We first show that if all the firms are connected, that is the graph $G$ has only one connected component, then $K = 1$. The $\text{rank}([X|B]) + \dim(\text{Null}([X|B])) = N + A$, where $\text{Null}(\cdot)$ denotes the null space of a matrix. It is then enough to show that $\dim(\text{Null}([X|B])) = 1$. We now show that $\text{Null}([X|B]) = \{[c]_u : R = c1_N \text{ and } u_a = -c1_A \text{ for any } c \in \mathbb{R}\}$.

Obviously any $[X|B][c]_u = 0$. Now consider $R$ and $u$ satisfying $XR + Bu = 0$. For any two firms $i \sim j$ such that $a \in \mathcal{A}_{ij}$ we have that $R_i = R_j = u_a$. Then $R_i = R_j$ whenever $i \sim j$.

Since $\mathcal{N}$ is connected, there is a path connecting any two firms, thus all $R_i$ are equal, and by the equation $XR + Bu$ all $u_a$ are equal to $u_a = -R_i$. Therefore, $\text{rank}([X|B]) = N + A - 1$.

Suppose now that there are $K$ connected components. We will show that this implies $\text{rank}([X|B]) = N + A - K$.

By item (ii), the set of firms $\mathcal{N}$ and the set of analysts $\mathcal{A}$ can be partitioned into $K$ connected (disjoint) components $\mathcal{N}^{(k)}$ and $\mathcal{A}^{(k)}$ such that $\mathcal{N} = \bigcup_{k=1}^{K} \mathcal{N}^{(k)}$, $\mathcal{A} = \bigcup_{k=1}^{K} \mathcal{A}^{(k)}$ and such that any two stocks $i$ and $j$ in $\mathcal{N}^{(k)}$ are connected by analysts belonging to $\mathcal{A}^{(k)}$. After a relabeling of the firms and the analysts, we have that

$$X = \begin{bmatrix} X^{(1)} & 0 & \cdots & 0 \\ 0 & X^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X^{(K)} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B^{(1)} & 0 & \cdots & 0 \\ 0 & B^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B^{(K)} \end{bmatrix}$$

and each $\mathcal{N}^{(k)}$ is a connected by the set of analysts in $\mathcal{A}^{(k)}$. Since the subset $\mathcal{N}^{(k)}$ is connected

31
this implies that \( \text{rank}([X^{(k)}|B^{(k)}]) = N^{(k)} + A^{(k)} - 1 \).

But due to the block structure of \([X|B]\), the \( \text{rank}([X|B]) = \sum_{k=1}^{K} \text{rank}([X^{(k)}|B^{(k)}]) = \sum_{k=1}^{K} (N^{(k)} + A^{(k)} - 1) = N + A - K. \)

QED.

**Proof of Proposition 6:** We need to show that \( \sum_{i \in N^{(k)}} \Delta \omega_i = 0 \) or that \( 1_k^T \Delta \omega = 0 \) for all \( y \)

The key step is to show that

\[
1_k^T \left( \left( \left( (\tau \Sigma)^{-1} + X'QX \right)^{-1} + (1 - \tau) \Sigma \right)^{-1} \left( (\tau \Sigma)^{-1} + X'QX \right)^{-1} \right) = \theta 1_k^T
\]

where \( \theta = (1 + (1 - \tau) \tau^{-1}) \). This is sufficient for the proof because

\[
1_k^T \Delta \omega = \frac{1}{\gamma} \left( \left( (\tau \Sigma)^{-1} + X'QX \right)^{-1} + (1 - \tau) \Sigma \right)^{-1} \left( (\tau \Sigma)^{-1} + X'QX \right)^{-1} X'Qy
\]

\[
= \theta 1_k^T X'Qy
\]

\[
1_k^T X' = 1_k^T
\]

\[
1_k^T Q = 0
\]
To prove the statement above consider:

\[ 1_k^T \left( (\tau \Sigma)^{-1} + X'QX \right)^{-1} (1 - \tau) \Sigma \right)^{-1} \left( (\tau \Sigma)^{-1} + X'QX \right)^{-1} = \theta 1_k^T \Leftrightarrow \]

\[ ((\tau \Sigma)^{-1} + X'QX)^{-1} ((\tau \Sigma)^{-1} + X'QX)^{-1} \left((1 - \tau) \Sigma \right)^{-1} 1_k = \theta 1_k \Leftrightarrow \]

\[ \left( (\tau \Sigma)^{-1} + X'QX \right)^{-1} (1 - \tau) \Sigma \right)^{-1} \left( (\tau \Sigma)^{-1} + X'QX \right) 1_k = \theta 1_k \Leftrightarrow \]

\[ ((\tau \Sigma) (\tau \Sigma)^{-1} + X'QX)^{-1} (1 - \tau) \Sigma \right)^{-1} 1_k + (1 - \tau) \tau^{-1} 1_k = \theta 1_k \Leftrightarrow \]

\[ (I + (\tau \Sigma) X'QX)^{-1} 1_k + (1 - \tau) \tau^{-1} 1_k = \theta 1_k \Leftrightarrow \]

\[ (I + (\tau \Sigma) X'QX)^{-1} 1_k = 1_k \Leftrightarrow (I + (\tau \Sigma) X'QX) 1_k = 1_k \]

QED.

**Proof of Proposition 7:** Let us compute $X'QX$. We can represent $X$, $Q$, and $B$ in block format as

\[
X = \begin{bmatrix}
X_1 \\
\vdots \\
X_a \\
\vdots \\
X_A
\end{bmatrix},
Q = \begin{bmatrix}
Q_1 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \cdots & \cdots & 0 \\
0 & \cdots & Q_a & \cdots & 0 \\
\vdots & \cdots & 0 & \ddots & 0 \\
0 & 0 & 0 & \cdots & Q_A
\end{bmatrix},
\text{and } B = \begin{bmatrix}
1_1 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \cdots & \cdots & 0 \\
0 & \cdots & 1_a & \cdots & 0 \\
\vdots & \cdots & 0 & \ddots & 0 \\
0 & 0 & 0 & \cdots & 1_A
\end{bmatrix}
\]

Thus $X'QX = \sum_{a=1}^A X_a'Q_aX_a$, where $Q_a = \Omega_a^{-1} - (1_a'\Omega_a^{-1}1_a)^{-1}(\Omega_a^{-1}1_a)(1_a'\Omega_a^{-1})$, $\Omega_a^{-1} = \text{diag}(\tau_{ia})_{i=1}^N_a$, and $1_a$ is the vector of $|N_a|$ ones. Note that $(1_a'\Omega_a^{-1}1_a)^{-1} = \left(\sum_{k \in N_a} \tau_{ka}\right)^{-1}$ and $[(\Omega_a^{-1}1_a)(1_a'\Omega_a^{-1})]_{ij} = \tau_{ia}\tau_{ja}$.

The element in row $ia$ and column $ja$ of matrix $Q$ is $-\tau_{ia}\tau_{ja} \left(\sum_{k \in N_a} \tau_{ka}\right)^{-1}$ for $i \neq j$ in
\[ N_a, \text{ the diagonal element is } \tau_{ia} - \tau_{ia}^2 \left( \sum_{k \in N_a} \tau_{ka} \right)^{-1}, \text{ and all other terms are zero. So the row and column sums of } Q \text{ are zero, and } Q \text{ has a positive diagonal and negative o-diagonal terms.} \]

Therefore, \[ [X'_a Q_a X_a]_{ij} = \tau_{ia} \tau_{ja} \left( \sum_{k \in N_a} \tau_{ka} \right)^{-1} \text{ for } i \neq j \] and \[ [X'_a Q_a X_a]_{ii} = \tau_{ia} \tau_{ja} \left( \sum_{k \in N_a} \tau_{ka} \right)^{-1} \text{ for } i \neq j. \] This implies that \( L(G) = X'QX. \) QED

**Proof of Lemma 3:**

(i) We have just shown that

\[ V = \text{var} (\theta|y) = \left( \tau^{-1} \Sigma^{-1} + X'QX \right)^{-1} \]

and that \( X'QX = D - A. \) Thus \( V = \left( (\tau \Sigma)^{-1} + D - A \right)^{-1}. \) Let diagonal matrix \( \Theta \) be defined by \( \Theta = (\tau \Sigma)^{-1} + D, \) so that \( V = (\Theta - A)^{-1} \) where \( \Theta = \text{diag}(\theta_i)_{i=1}^N \geq 0 \) and \( A \geq 0. \)

The matrix \( \Theta - A \) is a nonsingular \( M \)-matrix. Indeed, let \( s = \max_{i \in N} (\theta_i) > 0. \) Then \( \Theta - A = sI - B, \) where \( B = A + sI - \Theta \geq 0. \) The matrix \( \Theta - A \) is a nonsingular \( M \)-matrix because \( s > \rho(B), \) where \( \rho(B) \) is the spectrum radius of \( B. \) This holds because the matrix \( \Theta - A \) is strictly dominant diagonal,

\[ (\Theta - A)_{ii} = \left( \tau \sigma_{ia}^2 \right)^{-1} + d_i > d_i = \sum_j A(G)_{ij} = \sum_{j \neq i} |(\Theta - A)_{ij}| \]

and thus \( s > \rho(B). \)

Any nonsingular \( M \)-matrix is invertible and has a non-negative inverse, which proves that the inverse of \( (\Theta - A) \) exists and is nonnegative. In order to show that the inverse satisfy \( V_{ii} \geq V_{ij} \) for all \( i \) and \( j, \) we use the following lemma (Berman and Plemmons, 1994, pp. 254): Let \( W \) be a nonsingular \( M \)-matrix whose row sum are all nonnegative. Then \( V = (W)^{-1} \) satisfy \( V_{ii} \geq V_{ij} \) for all \( i \) and \( j. \)

(ii) We can normalize \( V \) multiplying by the diagonal matrix \( \Theta^{-1/2}, \)

\[ V = (\Theta - A)^{-1} \Theta^{-1/2} (I - \Theta^{-1/2} A \Theta^{-1/2})^{-1} \Theta^{-1/2} \]

34
But the spectrum radius of $\Theta^{-1/2}A\Theta^{-1/2}$, $\rho(\Theta^{-1/2}A\Theta^{-1/2}) < 1$ which implies that

$$(I - \Theta^{-1/2}A\Theta^{-1/2})^{-1} = \sum_{k=0}^{\infty} (\Theta^{-1/2}A\Theta^{-1/2})^k,$$

which completes the proof.

QED

**Proof of Proposition 8:**

The conditional returns $E(R|y) = ((\tau \Sigma)^{-1} + X'QX)^{-1}X'Qy = VX'Qy$ are a linear function of the stock recommendations so the derivative with respect to $y_{ia}$ is equal to

$$\frac{\partial E(R|y)}{\partial y_{ia}} = VX'Q_{(ia)}$$

where $Q_{(ia)}$ is the $ia$-th column of the matrix $Q$.

The column $Q_{(ia)}$ is as follows: the element in row $ja$ is $-\tau_{ia} \tau_{ja} (\sum_{k \in N_a} \tau_{ka})^{-1}$ for $j \neq i$ in $N_a$, is $\tau_{ia} - \tau_{ia} \tau_{ja} (\sum_{k \in N_a} \tau_{ka})^{-1}$ for $i = j$ in $N_a$, and is zero otherwise for $j \notin N_a$.

Let $\bar{\tau}_{ia} = \frac{\tau_{ia}}{\sum_{k \in N_a} \tau_{ka}}$, thus the vector $v_{ia} = X'Q_{(ia)}$ is simply $(v_{ia})_j = -\tau_{ia} \bar{\tau}_{ja}$ for $j \neq i$ in $N_a$, is $(v_{ia})_i = \tau_{ia} (1 - \bar{\tau}_{ia})$ for $i = j$ in $N_a$, and is zero otherwise. Therefore,

$$\frac{\partial E(R|y)_j}{\partial y_{ia}} = (Vv_{ia})_j = V_{ji} (\tau_{ia} - \tau_{ia} \bar{\tau}_{ia}) - \sum_{k \in N_a \setminus i} \tau_{ia} V_{jk} \bar{\tau}_{ka}$$

$$= \tau_{ia} \left( V_{ji} (1 - \bar{\tau}_{ia}) - \sum_{k \in N_a \setminus i} V_{jk} \bar{\tau}_{ka} \right)$$

$$= \tau_{ia} \left( V_{ji} - \sum_{k \in N_a} V_{jk} \bar{\tau}_{ka} \right) .$$

Therefore $\frac{\partial E(R|y)_j}{\partial y_{ia}} \geq 0$ if and only if $V_{ji} \geq \sum_{k \in N_a} V_{jk} \bar{\tau}_{ka}$. Also $\frac{\partial E(R|y)_i}{\partial y_{ia}} = \tau_{ia} (\sum_{k \in N_a} (V_{ii} - V_{ik}) \bar{\tau}_{ka}) \geq 0$ because $V_{ii} \geq V_{ik} \geq 0$, which completes the proof.

QED
Proof of Proposition 9:

(i) By the Sherman-Morrison-Woodbury formula

\[ \text{var}(R)^{-1} = \Sigma^{-1} = (\Sigma_{\varepsilon} + \beta'\beta)^{-1} = \Sigma_{\varepsilon}^{-1} - (1 + \beta'\Sigma_{\varepsilon}^{-1}\beta)^{-1}\Sigma_{\varepsilon}^{-1}\beta'\Sigma_{\varepsilon}^{-1}. \]

But \( \text{var}(\theta|y) = (\tau^{-1}\Sigma^{-1} + X'QX)^{-1} \), thus

\[ \text{var}(\theta|y) = \left(\tau^{-1}\Sigma_{\varepsilon}^{-1} + X'QX - \tau^{-1}(1 + \beta'\Sigma_{\varepsilon}^{-1}\beta)^{-1}\Sigma_{\varepsilon}^{-1}\beta'\Sigma_{\varepsilon}^{-1}\right)^{-1} \]

A second application of the Sherman-Morrison-Woodbury formula yields

\[ \text{var}(\theta|y) = \left(\tau^{-1}\Sigma_{\varepsilon}^{-1} + X'QX\right)^{-1} + \kappa \left(\tau^{-1}\Sigma_{\varepsilon}^{-1} + X'QX\right)^{-1}\Sigma_{\varepsilon}^{-1}\beta'\Sigma_{\varepsilon}^{-1} \left(\tau^{-1}\Sigma_{\varepsilon}^{-1} + X'QX\right)^{-1} \]

so that \( \text{var}(\theta|y) = V + \kappa b b' \), and \( \text{var}(R|y) = V + \kappa b b' + (1 - \tau)(\Sigma_{\varepsilon} + \beta'\beta) \), where

\[ \kappa = \left(\tau \left(1 + \beta'\Sigma_{\varepsilon}^{-1}\beta\right) - \beta'\Sigma_{\varepsilon}^{-1} \left(\tau^{-1}\Sigma_{\varepsilon}^{-1} + X'QX\right)^{-1}\Sigma_{\varepsilon}^{-1}\beta\right)^{-1}, \]

\[ b = \left(\tau^{-1}\Sigma_{\varepsilon}^{-1} + X'QX\right)^{-1}\Sigma_{\varepsilon}^{-1}\beta. \]

It remains to show that \( \kappa > 0 \), which is equivalent to,

\[ \tau \left(1 + \beta'\Sigma_{\varepsilon}^{-1}\beta\right) > \beta'\Sigma_{\varepsilon}^{-1} \left(\tau^{-1}\Sigma_{\varepsilon}^{-1} + X'QX\right)^{-1}\Sigma_{\varepsilon}^{-1}\beta. \]

We use the following lemma (see Horn and Johnson (2006, pp. 471): If A and B are positive definite matrices then \( A \succ B \) (i.e., \( A - B \) is positive semidefinite) if and only if \( A^{-1} \prec B^{-1} \).
The matrix $X'QX$ is positive semidefinite, so by the lemma above we have

$$
\tau^{-1}\Sigma^{-1}_\varepsilon + X'QX \succeq \tau^{-1}\Sigma^{-1}_\varepsilon
$$

$$
\tau \Sigma_\varepsilon \succeq (\tau^{-1}\Sigma^{-1}_\varepsilon + X'QX)^{-1}
$$

Thus

$$
\tau \beta' \Sigma^{-1}_\varepsilon \beta = (\beta' \Sigma^{-1}_\varepsilon) \tau \Sigma_\varepsilon (\Sigma^{-1}_\varepsilon \beta) \geq \beta' \Sigma^{-1}_\varepsilon \left( \tau^{-1}\Sigma^{-1}_\varepsilon + X'QX \right)^{-1} \Sigma^{-1}_\varepsilon \beta
$$

which implies that $\kappa > 0$ and concludes the proof of (i) since $\tau > 0$ so $\tau (1 + \beta' \Sigma^{-1}_\varepsilon \beta) > \tau \beta' \Sigma^{-1}_\varepsilon \beta$.

(ii) The conditional returns $E(R|y) = \left( (\tau \Sigma)^{-1} + X'QX \right)^{-1} X'Qy = (V + \kappa b b') X'Qy$ are a linear function of the stock recommendations so the derivative with respect to $y_{ia}$ is equal to

$$
\frac{\partial E(R|y)}{\partial y_{ia}} = (V + \kappa b b') X'Q_{(ia)}
$$

where $Q_{(ia)}$ is the $ia$-th column of the matrix $Q$, and the column $Q_{(ia)}$ is as defined in the proof of the previous proposition. Therefore,

$$
\frac{\partial E(R|y)}{\partial y_{ia}} = (V + \kappa b b') v_{ia},
$$

and the first part of the expression is

$$
(V v_{ia})_j = \tau_{ia} \left( V_{ji} - \sum_{k \in N_a} V_{jk} \tau_{ka} \right)
$$
and the second part is

\[(κbb'v_{ia})_j = κτ_{ia}b_jb_i (1 − \bar{τ}_{ia}) − \sum_{k \in \mathcal{N}_a \setminus i} κτ_{ia}b_jb_k \bar{τ}_{ka} \]

\[= κτ_{ia}b_j \left( b_i (1 − \bar{τ}_{ia}) − \sum_{k \in \mathcal{N}_a \setminus i} b_k \bar{τ}_{ka} \right) \]

\[= κτ_{ia}b_j \left( b_i − \sum_{k \in \mathcal{N}_a} b_k \bar{τ}_{ka} \right) \]

\[= κτ_{ia}b_j \sum_{k \in \mathcal{N}_a} (b_i − b_k) \bar{τ}_{ka}. \]

Therefore,

\[
\begin{align*}
\frac{∂E (R|y)}{∂y_{ia}} &= τ_{ia} \left( \left( V_{ji} − \sum_{k \in \mathcal{N}_a} V_{jk} \bar{τ}_{ka} \right) + κb_j \sum_{k \in \mathcal{N}_a} (b_i − b_k) \bar{τ}_{ka} \right) \\
&= τ_{ia} \left( \sum_{k \in \mathcal{N}_a} (V_{ji} − V_{jk}) \bar{τ}_{ka} + κb_j \sum_{k \in \mathcal{N}_a} (b_i − b_k) \bar{τ}_{ka} \right) \\
&= τ_{ia} \left( \sum_{k \in \mathcal{N}_a} (V_{ji} − V_{jk} + κb_j (b_i − b_k)) \bar{τ}_{ka} \right),
\end{align*}
\]

which completes the proof.

QED

**Proof of Corollary 3:** We show that when the betas are the same for all stocks, that is \(β = γ1\), then

\[b = (τ^{-1}Σ^{-1}_ε + X'QX)^{-1}Σ^{-1}_ε β = τβ.\]

This is sufficient to complete the proof since in the expression for the derivative

\[
\frac{∂E (R|y)}{∂y_{ia}} = τ_{ia} \left( \left( V_{ji} − \sum_{k \in \mathcal{N}_a} V_{jk} \bar{τ}_{ka} \right) + κb_j \sum_{k \in \mathcal{N}_a} (b_i − b_k) \bar{τ}_{ka} \right)
\]

the second term cancels out when \(b = τγ1\).
Indeed \( b = \tau \beta \) because

\[
(\tau^{-1}\Sigma^{-1}_\varepsilon + X'QX) b = \tau^{-1}\Sigma^{-1}_\varepsilon \tau \beta = \Sigma^{-1}_\varepsilon \beta
\]

because \( X'QX1 = 0 \). This completes the proof.

QED

**Proof of Corollary 4:** In order to explore the economics this relationship consider \( \Sigma \varepsilon = \sigma^2 I \) so that

\[
b = (\tau^{-1}\Sigma^{-1}_\varepsilon + X'QX)^{-1} \Sigma^{-1}_\varepsilon \beta = \sigma^{-2}_\varepsilon (\tau^{-1}\sigma^{-2}_\varepsilon I + X'QX)^{-1} \beta.
\]

Consider the eigendecomposition of \( X'QX \), \( \Lambda P \) where \( P = [v_1, ..., v_n] \) are the eigenvectors with eigenvalues \( \lambda_1 = 0 < \lambda_2 \leq \lambda_3 \leq ... \lambda_N \).

Let

\[
\beta = \gamma_1 v_1 + \gamma_2 v_2 + ...
\]

\[
\beta = \langle \beta, v_1 \rangle v_1 + \langle \beta, v_2 \rangle v_2 + ...
\]

Then we have that

\[
b = \sum_{k=1}^{n} \left( \frac{\sigma^{-2}_\varepsilon \langle \beta, v_k \rangle}{\tau^{-1}\sigma^{-2}_\varepsilon + \lambda_k} \right) v_k
\]

Given that \( \beta = c + dv_2 = \beta = cv_1 + dv_2 \) then

\[
b = \sum_{k=1}^{2} \left( \frac{\sigma^{-2}_\varepsilon \langle \beta, v_k \rangle}{\tau^{-1}\sigma^{-2}_\varepsilon + \lambda_k} \right) v_k = \theta 1 + \left( \frac{\sigma^{-2}_\varepsilon d}{\tau^{-1}\sigma^{-2}_\varepsilon + \lambda_2} \right) v_2
\]

\[
(b_i - b_k) = \left( \frac{\sigma^{-2}_\varepsilon d}{\tau^{-1}\sigma^{-2}_\varepsilon + \lambda_2} \right) ((v_2)_i - (v_2)_k) = \left( \frac{\sigma^{-2}_\varepsilon}{\tau^{-1}\sigma^{-2}_\varepsilon + \lambda_2} \right) (\beta_i - \beta_k)
\]
\[ b_i - \bar{b}_{N_a} = \sum_{k=2}^{n} \langle \beta, v_k \rangle \left( \frac{1}{\tau^{-1}\sigma_{\varepsilon}^{-2} + \lambda_k} \right) \left( (v_k)_i - \frac{1}{|N_a|} \sum_{m \in N_a} (v_k)_m \right) \]

The coefficient for the smallest eigenvalue \( \left( \frac{1}{\tau^{-1}\sigma_{\varepsilon}^{-2} + \lambda_2} \right) \geq \left( \frac{1}{\tau^{-1}\sigma_{\varepsilon}^{-2} + \lambda_k} \right) \)

Suppose that \( \langle \beta, v_2 \rangle \) is big and the other values are small \( \langle \beta, v_k \rangle = 0 \) for all \( k > 2 \).

We then have that

\[ b_i - \bar{b}_{N_a} = \left( \frac{\langle \beta, v_2 \rangle}{\tau^{-1}\sigma_{\varepsilon}^{-2} + \lambda_2} \right) \left( (v_2)_i - \frac{1}{|N_a|} \sum_{m \in N_a} (v_2)_m \right) \]

\[ b_i - \bar{b}_{N_a} = \left( \frac{\sigma_{\varepsilon}^{-2}}{\tau^{-1}\sigma_{\varepsilon}^{-2} + \lambda_2} \right) \left( \beta_i - \frac{1}{|N_a|} \sum_{m \in N_a} \beta_m \right) \]

\[ b_i - \bar{b}_{N_a} = \left( \frac{\sigma_{\varepsilon}^{-2}}{\tau^{-1}\sigma_{\varepsilon}^{-2} + \lambda_2} \right) \left( \beta_i - \bar{\beta}_{N_a} \right) \]

QED