Why Do Workers at Larger Firms Outperform?*

Rex Wang Renjie† Shuo Xia‡

June 12, 2020

ABSTRACT

Workers at larger firms outperform on average. For example, equity analysts working for more prestigious brokerage firms produce more accurate earnings forecasts. Analysts employed by larger brokerages are about 6% more accurate than analysts employed by small brokerages, which is equivalent to an advantage of 17.5 years of more experience. This finding is driven by two significant effects: more reputable firms provide more resources that improve analysts' forecasting ability (influence), and more reputable firms attract more talented candidates (sorting). We use a two-sided matching model to disentangle these two effects. We find that the direct influence effect accounts for 73% of the total impact while the sorting effect accounts for the remaining 27%.

*This draft is very preliminary. Please do not cite.

†Vrije Universiteit (VU) Amsterdam and Tinbergen Institute. E-mail: renjie-rex.wang@vu.nl.

‡Leipzig University and Halle Institute for Economics Research. E-mail: Shuo.Xia@iwh-halle.de.

1 Introduction

Workers at more prestigious companies have better performance on average. For example, academic researchers at higher ranked schools have better publication records; attorneys at larger law firms win more court cases; and sell-side equity analysts employed by more reputable brokerage houses produce more accurate earnings forecasts. This performance premium is driven by two distinct effects: the direct effect (influence) of more resourceful employers and the selection effect because of sorting in the labor market, which leads more prestigious companies to hire better candidates. This sorting mechanism creates an endogeneity problem, making it troublesome to establish and quantify the causal effect of employers on workers' performance.

The purpose of this paper is to disentangle those two confounding effects and quantify their relative importance, by estimating a two-sided matching model for the labor market of sell-side equity analysts. Analysts play an important role in gathering, analyzing, and distributing information in financial markets. Their most important outputs are earnings forecasts, and they have strong incentives to accurate predictions. Mikhail et al. (1999), Hong et al. (2000), and Groysberg et al. (2011) show that more accurate forecasts can help analysts avoid job termination or move down to less reputable brokerage firms, especially for early career analysts. Also, Stickel (1992) and Groysberg et al. (2011) show that analysts with greater forecast accuracy are more likely to be nominated as "All-star" analysts and earn higher compensation.

We find that new analysts working for more reputable brokerage firms are more accurate on average. An analyst employed by the most reputable brokerage is about 6% more accurate than an analyst employed by a minor brokerage, which is equivalent to an advantage of 17.5 years of more experience. This performance premium is driven by the fact that more reputable brokerage firms have more resources that improve analysts' forecast accuracy; and by the sorting effect, whereby more reputable brokerage firms attract more talented analysts who are intrinsically better forecasters. Using a two-sided matching model, we are able to quantify the relative importance of these two distinct effects in determining analyst forecast accuracy. We find that both effects are important, and the influence effect accounts for 73% of the total effect of brokerage firms' reputation on analyst forecast accuracy, while the sorting effect accounts for the remaining 27%.

More reputable brokerage houses can help their new analysts improve their forecast accuracy in several ways. First, analysts working for more reputable brokerage firms may have access to better data and research support (Clement, 1999). Better information acquisition and analysis in more reputable brokerage houses lead to more accurate forecast results. Second, analysts working for more reputable brokerage firms may have better personal communication opportunities with the management teams they follow (Clement, 1999), and private interactions with these teams is one of the most influential factors that determine forecast accuracy (Soltes, 2013; Brown et al., 2015). On the other hand, sorting captures the effect that better-talented analysts are attracted to work for more reputable brokerage firms. Therefore, even if brokerage firms' reputations have no direct impact on analysts' forecast performance, we still observe that analysts who work for higher-reputation brokerage firms perform better, because the sorting effect leads to positive assortative matching between analysts' individual talent and broker reputation.

Distinguishing these two effects is challenging. Brokerage firm reputation becomes endogenous when better-talented analysts work for more reputable firms, and analysts' talent cannot be perfectly measured. The unobserved part of talent can then be correlated with the brokerage firm reputation measure, and the estimated effect of brokerage firm reputation will be biased upward. This concern increases when we focus on new analysts where the datasets contain little information on their abilities. The ideal solution to this endogeneity problem is to find an instrumental variable that is independent of an analyst forecasting ability but correlates with the reputation of the brokerage house hiring this analyst. However, the matching decision between analysts and brokerage firms are mutual choices, and it is a complicated process involving a number of observable and unobservable factors. To the best of our knowledge, there is no valid instrument that solve this endogeneity problem.

To circumvent this endogeneity issue, we take a structural approach similar to Sorensen (2007a). Our structural model contains two key elements: first, an outcome equation that models the determinants of analysts forecast accuracy, and second, a one-to-many associative matching model that captures the sorting process. The matching model explicitly models the matching process between analysts and brokerage firms and allows for matching decisions to interact with different agents. The matching decision interaction between agents creates difficulties in estimating the model, but also provides a rank order property that is useful for identification. The rank order property of the two-sided matching model means that the matching decision depends on the relative ranking of the agents in the market. Therefore, it not only depends on the characteristics of the matched agents themselves, but also on the other agents' characteristics. If the agents' characteristics vary exogenously across the market, we can identify the sorting effect by comparing the performance difference between analysts of different quality but match with brokerage firms with similar reputations in different markets. Similarly, we can identify the influence effect by comparing the performance difference between analysts with similar quality but match with brokerage firms with different reputation in different markets.

The key identification assumption is that agents are exogenously assigned across different markets. That is, we need sufficient variation across the new analyst labour market, and agents cannot choose to participate in a particular market for reasons correlated with the agents' characteristics in that market. In this study, we assume the new analyst labour market is segregated by the calendar year and geographically. A similar identification assumption has been made in Sorensen (2007a), Park (2013), Chen (2014), Ni and Srinivasan (2015), Pan (2015), Akkus et al. (2016a), and Xia (2018).

Agents' matching decisions interact, so any analysis of the likelihood function of one agent's decision must also take account of other agents' decisions. The likelihood function then becomes a high dimensional integral function and it cannot be factored out, as in the standard Heckman selection model (Heckman, 1979) where agents' decisions are independent of each other. To overcome the numerical difficulty in solving the high dimensional integration problem, we apply a Bayesian approach, use the Markov Chain Monte Carlo (MCMC) method to transform the integration problem into a simulation problem to make estimation feasible (Tanner and Wong, 1987; Albert and Chib, 1993; Sorensen, 2007a; Park, 2013; Chen, 2014; Ni and Srinivasan, 2015).

Our research contributes first to the literature on the determinants of analyst forecast accuracy. Brokerage firm resources have been found to affect analyst forecast accuracy (Clement, 1999; Kothari et al., 2016), and because of the lack of an identification strategy the sorting effect cannot be disentangled from the total impact. Therefore, the influence on analyst forecast accuracy is unknown. Our results not only provide the first quantitative estimates of the influence effect of the brokerage firm but also quantify the relative importance of the influence and the sorting effects.

Second, our study contributes to the literature that uses the two-sided matching model to understand the incentives for agents to match and the outcomes of the matching results in markets such as the venture capital market (Sorensen, 2007a; Akkus et al., 2016a; Fox et al., 2018), the labour market (Agarwal, 2015; Pan, 2015; Matveyev, 2016; Xia, 2018), M&A market (Park, 2013; Akkus et al., 2016b), and the bank lending market (Chen and Song, 2013; Schwert, 2018).

The results of our study also help to understand workers' incentives to work for firms with good reputations, and the incentives for firms to maintain their reputations. Edmans (2011) finds that firms with better reputations on average perform better, and our results suggest that the reputation of a firm can serve as a sorting mechanism to attract talented employees, which is beneficial for firm performance. More talented employees also like to work for firms with good reputations, because they can scale their ability by using the firms' resources and achieve better personal performance and better future career outcomes. Our results suggest that for new analysts the influencing effect of firms' reputations is 2.7 times larger than the sorting effect. Therefore, the benefit of working for high-reputation firms is particularly attractive for new workers.

The remainder of the paper is organised as follows. In Section 2 the data and the OLS estimation results are discussed. Section 3 presents the theoretical and empirical model and a discussion of identification. Section 4 provides the estimation results. Section 5 concludes the paper.

2 Data and OLS results

2.1 Sample selection and key variables construction

We consider new hires by brokerage firms in each year between 1996 and 2013. Our data comes from the Institutional Brokers Estimate System (I/B/E/S) database, which collects analysts' earnings forecasts and recommendations for companies worldwide. We use the I/B/E/S Detail Recommendations File to identify the brokerage firm an analyst is employed by in any given year. The recommendation file starts in 1992 and expands its coverage over the first three years, so we only consider analysts who started in 1996 or later. We classify an analyst as a new hire in a given year if she appears for the first time in the dataset in that year, and stays at least for the subsequent four years in the dataset and works for the same brokerage firm. We cross-check with the I/B/E/S Detail Earnings History File to further exclude analysts who had previously issued any earnings forecasts, and those who do not issue any earnings forecasts at all. We manually search for the location of the brokerage firms and remove analysts employed by foreign broker houses that do not have any offices in the U.S. Our final sample consists of 1,815 analysts hired by 284 brokerage firms for the period between 1996 and 2013.

Figure 1 illustrates the geographic distribution of brokerage houses from our sample. We plot the number of firms in each state. A clear geographic clustering on the demand side can be clearly seen in the Northeastern states such as NY and MA, accounting for roughly 65% of our sample. We therefore divide the analysts into 36 markets: Northeastern states and the remaining states for 18 years from 1996 to 2013. Note that pooling the other states together into one labour market each year is less of a concern under the assumption that those small local markets are independent of each other.[1]

To measure an analyst's performance, we first determine her accuracy for each stock she covers in a given year and then take the average of this accuracy across all coverage stocks over the first five years of her tenure. Specifically, for analyst i making a forecast for the earnings of fiscal year t of stock j, we compare her absolute forecast error to the

[1] We also run our analysis by considering all of the states as one big market in each year, and the results are similar.

average absolute forecast error of other analysts covering the same stock during the same time period. We rank all available absolute forecast errors from small to large and assign a rank that corresponds to the relative ranking of analyst i's forecast error for that stock-year. The analyst ranked n-th (where the most accurate/smallest error is ranked 1st and the least accurate/largest error is ranked Nth) is assigned.

$$rank_{ijt} = 1 - \frac{n_{ijt}}{N_{jt} + 1}. \tag{1}$$

The lower the rank, the less accurate the forecast. We aggregate those accuracy ranks for analyst i to determine her overall accuracy as

$$Accuracy_i = \frac{1}{5} \sum_{t=\tau}^{\tau+4} \left(\frac{1}{|J_{it}|} \sum_{j \in J_{it}} rank_{ijt} \right),$$

where J_t denotes analyst i's coverage in year t.

The brokerage firm prestige is measured by using Carter and Manaster (CM) ranking. This ranking measure is based on the order of brokerage firms in firms' IPO tombstone announcements. The measure is developed by Carter and Manaster (1990) and extended by Carter et al. (1998) and Loughran and Ritter (2004). We obtain the data from Jay Ritter's website. On a scale of 0 to 9, the higher the rank, the more prestigious the brokerage firms. Morgan Stanley, Goldman Sachs, JPMorgan, Deutsche Bank, and CITI Group are among the most frequently listed in the highest reputable brokerage groups.

Table I presents the summary statistics of our variables. The mean growth rate for these brokerage firms is 14.5% yearly, and the median growth rate is 5%. These firms are on average expanding through the sample period. The newly hired analysts on average start by covering slightly more than 8 stocks, less than the average number of stocks covered by analysts in the whole I/B/E/S universe, which is 14. Most of the analysts cover less than three different industries. The financial analyst labour market is racially dominated by white analysts, based on the surname search, and in our sample we classify less than 17% as nonwhite analysts. Analysts do not cluster in the main industries they cover in our sample. The largest group of analysts (27.9% of the total sample) cover firms in the high-tech industry, followed by 26.8%

who mainly cover industries other than those listed in the table. As over half of the U.S. publicly listed firms from 1996 to 2013 are classified in the high-tech industry or in "other" industries, this is a reasonable assumption.

2.2 Naive OLS results

In this subsection, we document a robust and strong empirical correlation between brokerage reputation and newly hired analysts' forecast performances. According to the level of brokerage prestige, we first plot the correlation between brokerage prestige and analyst performance.

Figure 2 illustrates strong positive correlations between broker prestige and analysts' accuracy and their likelihood of becoming an all-star analyst. Analysts who start with the lowest prestige brokerage firms on average exhibited performance of 0.493, while those who start with the highest prestige firms on average exhibited performance of 0.522, and those analysts are on average 6% more accurate. This effect is close to those documented in the literature (e.g. Clement (1999)), which is equivalent to the advantage of 17.5 years of more experience.

To investigate these relations more formally, we estimate an OLS model for analyst accuracy. Table II shows that for the entire 1996 - 2013 period, analysts work for higher prestige brokerage firms on average have greater forecast accuracy. The magnitude does vary when we include other broker and analyst characteristics in column (2) and market fixed effect in column (3). In column (4) and column (5) we repeat the analysis on subsamples from 1996 - 2004 and 2005 - 2013. Here, broker reputation is also positively correlated with analyst forecast accuracy. Overall, the positive correlation between broker prestige and analyst forecast accuracy is robust to different controls and split sample regressions. If an analyst moves from the lowest to the highest reputable group of brokerage firms, the analyst forecast accuracy will increase by 4.7% ($= 0.0234 \times 9/0.493$), which is equivalent to 13.8 years of more experience.

In addition to broker prestige, other factors affect newly hired analyst forecast accuracy. From Table II, we observe that the more stocks analysts cover, the more accurate their forecasts are. This observation may appear to contradict previous findings that the more complex the portfolios that analysts are covering, the less accurate their forecasts are. We argue this is less of a concern because our sample only contains newly hired analysts, so the number of stocks analysts cover also contains information on analysts' ability. Another critical factor explaining analyst forecast accuracy is the ethnicity of the analyst. In the whole sample, non-white analysts constitute less than 17% of the total sample but on average they perform better than white analysts. This outperformance is particularly strong in the first half of the sample, possibly because sell-side analyst jobs used to be occupied by white candidates and so the entry bar is higher for non-white candidates. For non-white candidates to get a job, their ability must be better than average, and thus they perform better[2].

As we explain in the introduction, the quality of brokerage firms becomes endogenous when sorting and causes more reputable brokerage houses to employ analysts who are better, along with many dimensions unobserved in the data. Analysts with better unobserved characteristics, as captured by the error term in the regression, match with brokers of better quality. The error term becomes positively correlated with broker size and broker accuracy, and the coefficient estimates are biased upwards relative to the brokers' actual influence. As no obvious instrumental variable is independent of analyst outcome but is related to the quality of the brokerage firm employing this analyst, we adopt the structural model developed by Sorensen (2007a) that exploits the implications of sorting to separate sorting from influence. Sorting implies that in a market with better broker firms, a given firm is pushed down the relative ranking and is left with worse analysts. Hence, a broker's new hire decisions depend on the characteristics of other agents in the market. Nevertheless, the outcome of the analyst is independent of these other characteristics, and the other brokers' characteristics serve as a source of exogenous variation. We now discuss the model in more detail.

[2]Similar evidence has been found in the asset management industry, where the entry bar is higher for candidates with low-income family backgrounds. Consequently, to become fund managers these candidates need to be significantly better than those from wealthy families (Chuprinin and Sosyura, 2018)

3 Model

3.1 Two-sided matching model

We model the labor market of sell-side analysts as a one-to-many two-sided matching market, which is based on the college admission model developed by Gale and Shapley (1962) and Roth and Sotomayor (1992) and is similar to the VC-entrepreneur matching model in Sorensen (2007a). Each firm can hire multiple analysts, while each analyst candidate can only be employed by one firm. However, in any given market, brokerage firms are restricted to the number of new analysts they can hire, as firms' hiring capacity is capped because of the limited demands and resources. Each potential match has a valuation (V), which represents the discounted expected future payoff of the possible matched pair. The brokerage firm receives λ fraction of the valuation, and the analyst expects to receive $1 - \lambda$ fraction, where λ is fixed for all possible matches in a market. Such setting rules out transfers and guarantees a unique equilibrium for the model. This assumption is reasonable because analysts are sharing profits of the firm. Even though we do not observe analyst compensation in general, most compensation is paid in the form of a bonus, which is high when a firm's bonus pool expands and low when it shrinks (Groysberg et al., 2011). In addition, because we focus on newly hired analysts, who have little bargaining power at the beginning of their career, it is unlikely that these analysts can negotiate on pay. Therefore, their compensation structure is mostly fixed, and they cannot match more reputable firms by being offered a lower profit share by the firm.

3.1.1 Agents

The matching model has two types of agents: analyst candidates and brokerage firms. In each market m, a set I_m contains all of the analyst candidates, and a set J_m contains all brokerage firms that are looking for new analysts. Each candidate will be employed by by one brokerage firm, and each brokerage firm can hire a limited number of analysts. Let brokerage firm j's quota be q_j, where $q_j > 0$. The set M_m contains all possible matches of analysts and firms in market m, therefore $M_m = I_m \times J_m$. A matching contains observed hirings in market m denoted as μ_m, where $\mu_m \subset M_m$. Denoting that μ_j contains all of the analysts firm i hires and μ_i is the brokerage firm analyst i works for, then a match between firm i and analyst j can be expressed as: $(i, j) \in \mu$, $i = \mu(j)$, or $j \in \mu(i)$.

Agents on both sides of the market choose their matched partners to maximise the matching value, which represents the expected latent joint utility at the time of hiring. Let each possible match have a matching value and let the value of the match i, j be denoted as $V_{i,j}$ regardless of whether i, j is a matched pair or not. The matching values are assumed to be distinct to avoid the possibility that agents can be indifferent between two matches. The matching utility is divided between the brokerage firms and analysts. Firms receive λ share of the matching value and the analysts receive $(1 - \lambda)$ share, and λ is fixed for all matches and $\lambda \in (0, 1)$.

3.1.2 Equilibrium

A matching is an equilibrium if it is stable and no pair of agents would like to deviate from their current matches and form a new match together to become a blocking pair. The stable equilibrium always exists (Gale and Shapley, 1962) and under the fixed sharing rule of the matching value the equilibrium is unique (Sorensen, 2007a). The unique equilibrium is characterised by a set of inequalities based on the no blocking pairs condition.

For i, j to be a stable match, we need no blocking pair to exist for i, j, that is, the opportunity cost of analyst i remaining match with firm j or the opportunity cost of firm j remaining match with analyst i has to be smaller than the matching value of i, j, $V_{i,j}$.

The opportunity cost of analyst i is the maximum value that analyst i can get from the feasible set of deviations of analyst i instead of working for the firm j. The opportunity cost of brokerage firm j is the maximum value that firm j can get from the feasible set of deviations of firm j instead of hiring analyst i. The fixed sharing rule means that finding the maximum value that agents on one side of the market can get is equivalent to find the maximum matching value that a pair of agents can achieve together. We denote OC_i as the corresponding matching value for analyst i's opportunity cost and OC_i is the corresponding matching value for brokerage firm j's opportunity cost. That is,

$$V_{i,j} < max[OC_i, OC_j],$$

where

$$OC_i \equiv max[V_{i,j'}], \forall j' \in J \cap (V_{i,j'} > V_{\mu(j'),j'}),$$
$$OC_j \equiv max[V_{i',j}], \forall i' \in I \cap (V_{i',j} > \min_{i'' \in \mu(j)} V_{i'',j}).$$

If in other circumstances analyst i and brokerage firm j are not matched, then (i, j) cannot become the blocking pair for their current matches. Then it is sufficient that,

$$V_{i,j} > max[V_{i,\mu(i)}, \min_{i''' \in \mu(j)} V_{i''',j}].$$

We denote $\overline{V}_{i,j} \equiv max[OC_i, OC_j]$, and $\underline{V}_{i,j} \equiv max[V_{i,\mu(i)}, \min_{i''' \in \mu(j)} V_{i'''j}]$. For μ to be a stable matching, the following conditions need to hold:

$$V_{i,j} < \overline{V}_{i,j}, \forall (i, j) \notin \mu, \tag{2}$$

$$V_{i,j} > \underline{V}_{i,j}, \forall (i, j) \in \mu. \tag{3}$$

3.2 Empirical Model

The first part of the empirical model is a matching function determining the matching value of the match between two agents. The matching value is unobserved and modelled as a latent variable. Without loss of generality, the matching value of analyst i and brokerage firm j can be written as:

$$V_{i,j} = \alpha W_{i,j} + \eta_{i,j}, \forall (i, j) \in M, \tag{4}$$

where $W_{i,j}$ contains characteristics of analyst i and firm j that are observed by econometricians. $\eta_{i,j}$ contains characteristics of analyst i and firm j that are not observed by econometricians but are known for every agent in the market and $\eta_{i,j} \sim N(0, \sigma_\eta)$.

The second part of the model is the outcome equation. This determines the outcome of all possible matches, which is only observable to those matches that are realised. The outcome of analyst i and brokerage firm j can be written as:

$$Y_{i,j} = \alpha X_{i,j} + \varepsilon_{i,j}, \forall (i,j) \in M, \tag{5}$$

where $X_{i,j}$ contains characteristics of analyst i and firm j that are observed by econometricians. $\varepsilon_{i,j}$ contains characteristics of analyst i and firm j that are not observed by econometricians but known for every agent in the market and $\varepsilon_{i,j} \sim N(0, \sigma_\varepsilon)$.[3]

Directly estimating the outcome equation leads to biased results, as the matching decision between analyst i and firm j is not random but correlated with the error term in the outcome equation, which cannot be observed by econometricians. This problem is captured by a third equation determining the correlation between the error terms in the valuation equation and the outcome equation:

$$\varepsilon_{i,j} = \delta \eta_{i,j} + \xi_{i,j}, \tag{6}$$

where $\xi_{ij} \sim N(0, \sigma_\xi)$. If there is no correlation between the two error terms then $\delta = 0$.

3.3 Identification and estimation

We now discuss how we identify and estimate the parameters in the outcome equation. The main feature of the matching market is that the agents' decisions on matching interact with each other, and this leads to better-talented analysts sorting by brokerage quality. If analyst A is hired by brokerage firm 1, then brokerage firm 2 cannot approach analyst A, as analyst A is not available anymore. Similarly, if brokerage firm 1 has used up its hiring quota, then other analysts with relatively lower quality than analyst A cannot match with broker 1 anymore. As such, in each market, agents' matching decisions do not only depend on their own qualities, but also correlate with other agents' characteristics.

The sorting and interaction feature helps us identify the direct influence effect from brokerage firms. As we rank all of the new analysts and all brokerage firms based on their

[3] If the outcome is binary, there will be a third part containing a binary outcome function, i.e. $O_{i,j} = 1[Y_{i,j} > 0]$

characteristics in each market, with the top-ranked analyst candidate matched with the top-ranked brokerage firm, we continue to match the second highest ranked analyst candidate with the top-ranked brokerage firm until the hiring quota is entirely filled, and then we continue to form matches between analysts with the second highest ranked brokerage firm until we fill all of the vacancies in the market. This rank-order property means the matching decision is determined by the relative ranking of the agents on two sides of the market, and partly depends on the agents' own characteristics, and partly on the characteristics of other agents. As the characteristics and quality of "other" agents vary between markets, similar-quality analysts would be matched with brokerage firms with different reputations for exogenous reasons, and can help to identify the parameters in the outcome equation.

The cross-market variation means that same-quality brokerages and same-quality analysts cannot match in two different markets. Assume in market 1, brokerage i and analyst j are matched. In market 2, brokerage i' has the same quality as brokerage i, but assume market 2 contains similar brokerage firms but with more talented analysts. Therefore, an analyst j' with the same quality as analyst j will rank much lower in market 2, and cannot match with brokerage i', and instead is matched with another brokerage firm with lower quality. Brokerage house i' can match with another analyst k who has better quality than analyst j'. The effect from matching is different, but the impact from the brokerage firm influence is the same, and this will lead to differences between outputs from analyst j and analyst k, which help us identify the effect of matching.

More formally, let Y^*_{ij} denote the observed match (i, j)'s outcome in one market, and then to estimate the coefficients based on the empirical model we have:

$$
\begin{aligned}
E[Y_{i,j}|X_{i,j}] &= E[Y^*_{i,j}|X_{i,j}, (i,j) \in \mu] \\
&= E[Y^*_{i,j}|X_{i,j}, V_{i,j} > \underline{V}_{i,j}] \\
&= \beta + E[\varepsilon_{i,j}|\alpha W_{i,j} + \eta_{i,j} > \underline{V}_{i,j}] \\
&= \beta + E[\delta\eta_{i,j} + \xi_{i,j}|\eta_{i,j} > \underline{V}_{i,j} - \alpha W_{i,j}] \\
&= \beta + \delta E[\eta_{i,j}|\eta_{i,j} > \underline{V}_{i,j} - \alpha W_{i,j}].
\end{aligned}
$$

The first equality comes from the equilibrium condition of the matching model, and the fourth equality comes from the error term correlation structure. Therefore, the exogenous variation in this expression identifies outcome equation parameters β, and the expression varies with $\underline{V}_{i,j}$. As $\underline{V}_{i,j}$ is determined by the other agents' characteristics in the market, if the allocation of the other agents in the market is exogenously given, then the parameters in the outcome equation are identified. [4]

The key identification assumption is that agents are allocated exogenously across markets, which is reasonable because the new analyst labour market is likely to be influenced by macro or financial industry factors instead of agents' sort on different markets (i.e., waiting to hire later because they know there will be better candidates one year later). Figure 3 shows that even though the average is reasonably consistent across markets, there are significant variations of main variables within each market, and this variation fluctuates from market to market. Thus, it is reasonable to assume the agents are exogenously allocated across markets.

The estimation method we use is the Bayesian estimation with Markov Chain Monte Carlo (MCMC). The sorting and interaction feature of the model makes estimation difficult. The likelihood function for one pair of agents' matching decisions also depends on the other agents' choices, so all of the error terms must be integrated simultaneously. To circumvent this high-dimensional integration problem, we take advantage of the Bayesian method with MCMC (Tanner and Wong, 1987; Geweke et al., 1994; Albert and Chib, 1993), and instead of solving the integration problem, we augment the observed data with the simulated value of the latent matching value and the performance of the counterfactual matches. The simulated distribution converges to the augmented posterior distribution. The detailed simulation procedure can be found in Appendix A.

[4] A complete discussion of the identification strategy can be found in Sorensen (2007b).

4 Estimation results

4.1 Main result

In this section, we estimate the structural model. In Table III Panel B, the coefficients estimated represent agents' preferences. The results show that an analyst prefers to work for brokerage firms with higher reputations and higher growth rates, and brokerage firms prefer to hire analysts who can cover large value portfolios, cover fewer stocks, have less industry focus, and are from a non-white background. Thus, firms prefer non-white analysts who can cover a limited amount of large firms and span less industries. The probability of an analyst match with a broker with the highest reputation score is 90.6%. The probability that brokerage firms prefer a non-white analyst is 55.13%. Compared with a new analyst who can only cover the lowest ten percentile of the portfolio market size, brokerage firms prefer analysts who cover the top ten percentile of the portfolio market size by a probability of 59.5%. Overall, the results from the matching equation suggest analysts have strong preferences in terms of broker reputation, rather than other observed factors that brokerage firms have on analysts. Broker reputation is the most important factor in measuring brokerage firms' quality, while the analysts' ethnicity or portfolio sizes are simply indirect measures of their quality.

Panel C of Table III represents the effect of sorting on unobserved characteristics. If there is no sorting between unobservables, a matching model is not needed. The result shows δ is positive and 0 is not contained in the 99% highest posterior distribution, and that the sorting effect exists and is significant, indicating that unobserved agents' characteristics affect matching values and also matching outcomes. This also highlights the key point of the study: controlling for matching is crucial given its significant effect.

Panel A of Table III shows the estimated coefficients in the outcome equation after controlling for endogenous matching. The coefficient associated with broker reputation is positive and 0 is not contained in the 95% highest posterior distribution, which suggests after controlling for sorting, the effect of brokerage reputation is crucial in explaining analyst forecast accuracy. This finding is consistent with channels suggested by Clement (1999) that brokerage resources (proxied by brokerage reputation in this study) are important in determining analyst forecast accuracy.

4.2 Relative importance

Although the above analysis clearly shows that broker reputation has a significant direct impact on analyst forecast accuracy because sorting on unobservables also has a significant impact on the outcome, the relative importance of the direct effect of broker reputation, and the indirect effect from sorting is unknown.

In determining the relative importance, we compare the OLS and Bayesian estimated results in Table IV. Column (1) presents the OLS regression results, and column (2) the Bayesian estimation results. Figure 4 shows how we decompose the total effect into the influence effect and the sorting effect. Controlling for sorting effect, analysts employed by the brokerage houses with the highest reputation rank are on average 3.5% ($= 0.0019 \times 9/0.493$) more accurate than those employed by the small brokers. This advantage in accuracy is comparable to 10.2 years more experience. On the other hand, the difference between the OLS and MCMC coefficient estimates of broker reputation indicates the selection effect because of sorting in the labor market, which is both economically and statistically significant. The selection effect accounts for 27% of the total effect estimated by the naive OLS regression, while the influence of brokers accounts for 73% of the total impact. oo

4.3 Alternative market

In our main analysis, our definition of new analyst labour market is by one calendar year but segregated by geographical locations. The market segregation is a critical identification assumption, and will fail if new analyst candidates or brokerage firms choose to participate in the specific market, based on unobserved characteristics of other agents in that market. For example, if the Northeast of the US has more reputable brokerage firms and if that reputation is sufficient to attract analyst candidates, this will lead to analysts sorting between different locations, and so a more appropriate definition of the market is to consider the whole US as a single market.

In this subsection, we expand the market definition to evaluate the robustness of the estimation results. In Column (3) of Table IV, we treat the Northeast and the rest of the US as the same market and repeat the analysis. The estimated coefficients are at a similar magnitude and significance level, particularly the key variables of broker reputation and δ. The magnitude of the coefficient associated with broker reputation is robust to different specifications and the statistical significance is also similar. The magnitude of δ increases but the statistical significance is similar. This indicates that minor sorting exists between the geographical locations in the same year, but the baseline Bayesian estimation does not capture this minor effect. For our main purpose of estimating the direct effect of broker reputation, this is less of a concern because this cross-location sorting appears to have little correlation with broker reputation. Overall, the results provide an intuitive robustness test that confirms that the identification assumption is valid and our estimation results are not sensitive to different market definitions.

5 Conclusion

Our study focuses on the new analyst labor market. We find new analysts working for firms with higher reputations perform better. This total effect is a combination of the direct influence effect, in which reputable firms can help analysts perform better, and the sorting effect, in which brokerage firms with high reputations can attract more talented analysts. To disentangle these two effects, we utilize a one-to-many two-sided matching model to circumvent the need to find the instrumental variable. The features of the matching model can capture how agents' matching decisions interact, and how the other agents' characteristics determine the relative ranking of the agents' matching decisions, but the other agents' characteristics do not have an effect on the agents' performance. Therefore, the exogenous variation of the other agents' characteristics helps to identify the coefficients of the outcome equation.

In the sample of 1815 new analyst-brokerage firm matched pairs from 1996 to 2013, we find that both the influence effect and the sorting effect have a significant impact on analyst forecast accuracy. The influence effect accounts for 73% of the total impact, and the sorting effect for 27%.

The results of the study have more general implications for understanding the incentives for workers to choose more reputable firms to work for and the incentives for firms to spend resources in maintaining their reputation. High reputation firms provide resources for workers and help them perform better, and in our results the forecast difference between analysts of the same quality working for the lowest and the highest reputation firms is equivalent to 15 years of experience. A firm's reputation is valuable, as it not only motivates current workers but also attracts more talented new workers. Both of these effects are important in understanding the benefit of firms' reputation on workers performances.

Figure 1: Geographic distribution of brokerage firms

This figure shows the distribution of the US brokerage firms' headquarters in different states. The darker the states, the more brokerage firms are located in that state.

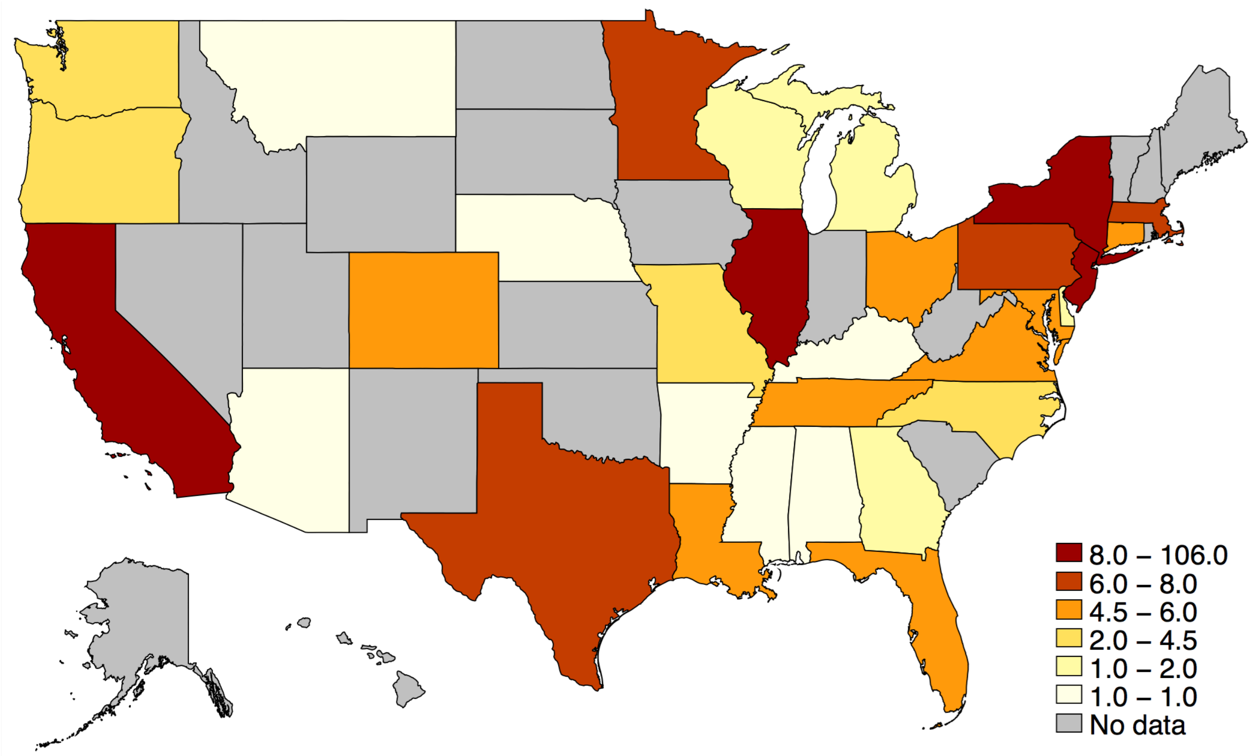


Figure 2: Relation between brokerage firm prestige and analyst performance

This figure shows the correlation between brokerage firms' reputations and newly hired analysts' forecast accuracy from 1996 to 2013. Our sample is grouped into 10 bins according to broker prestige. The shadow area represents a 95% confidence interval.

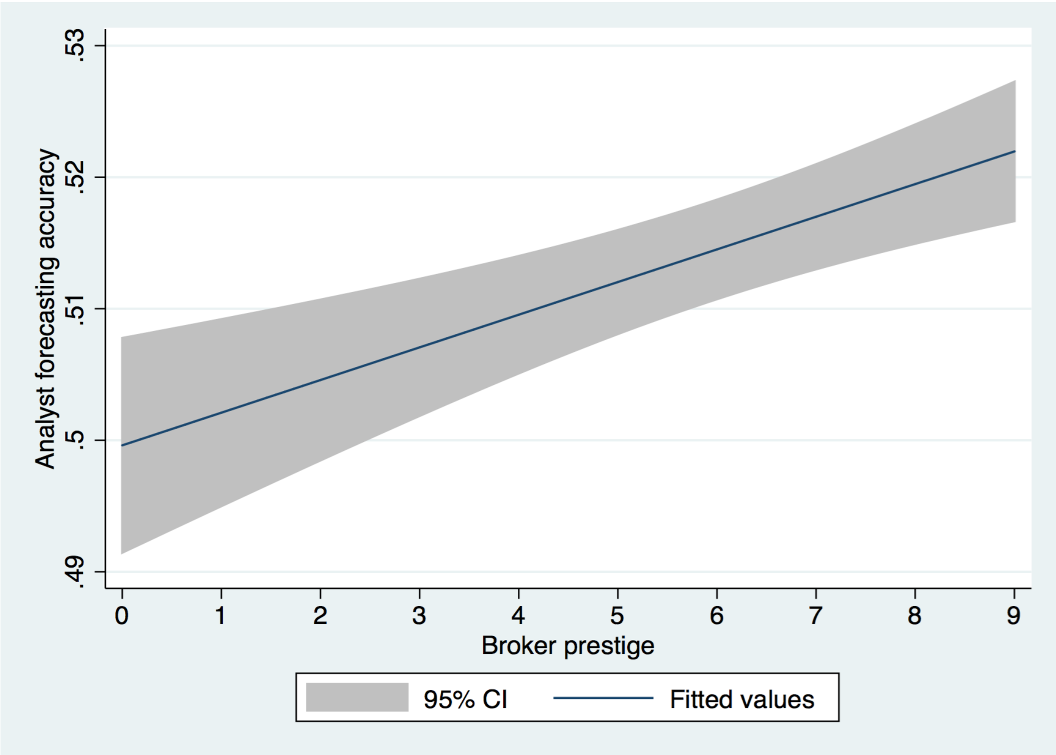
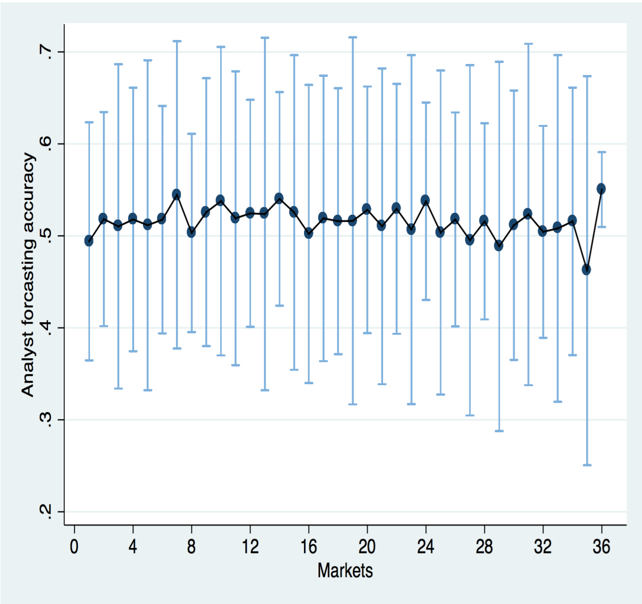
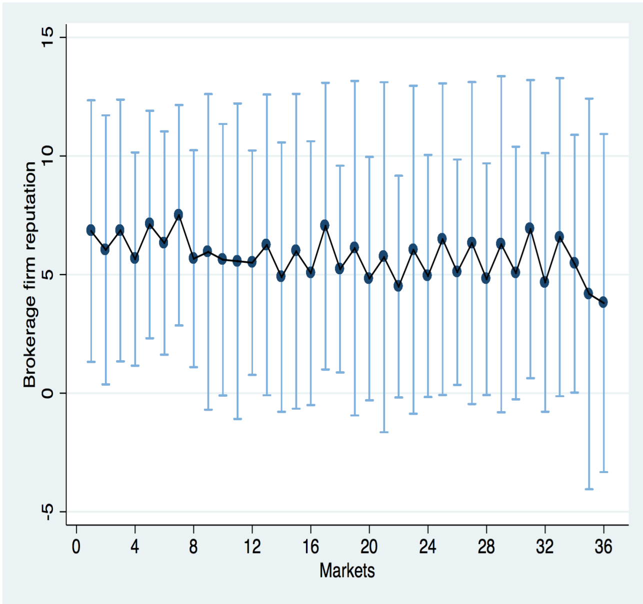


Figure 3: Variation in main variables across markets

This figure shows the cross and within market variation of the key Y variable: analyst forecast accuracy, and the key X variable: broker reputation. Each subgraph depicts the average of the variable (black solid line) and one standard deviation around the mean (light blue error bar). Subgraph (a) shows the variation of analyst accuracy across different markets. Subgraph (b) shows the variation of brokerage reputation across different markets.



(a) Analyst forecast accuracy



(b) Broker reputation

Figure 4: Decomposition of influence and sorting

This figure shows the results comparison between the naive OLS regression results and the sorting controlled outcome equation results in Table IV for the highest reputable brokerage firms. The dashed line represents influence effect of the broker reputation after controlling for the selection effect.

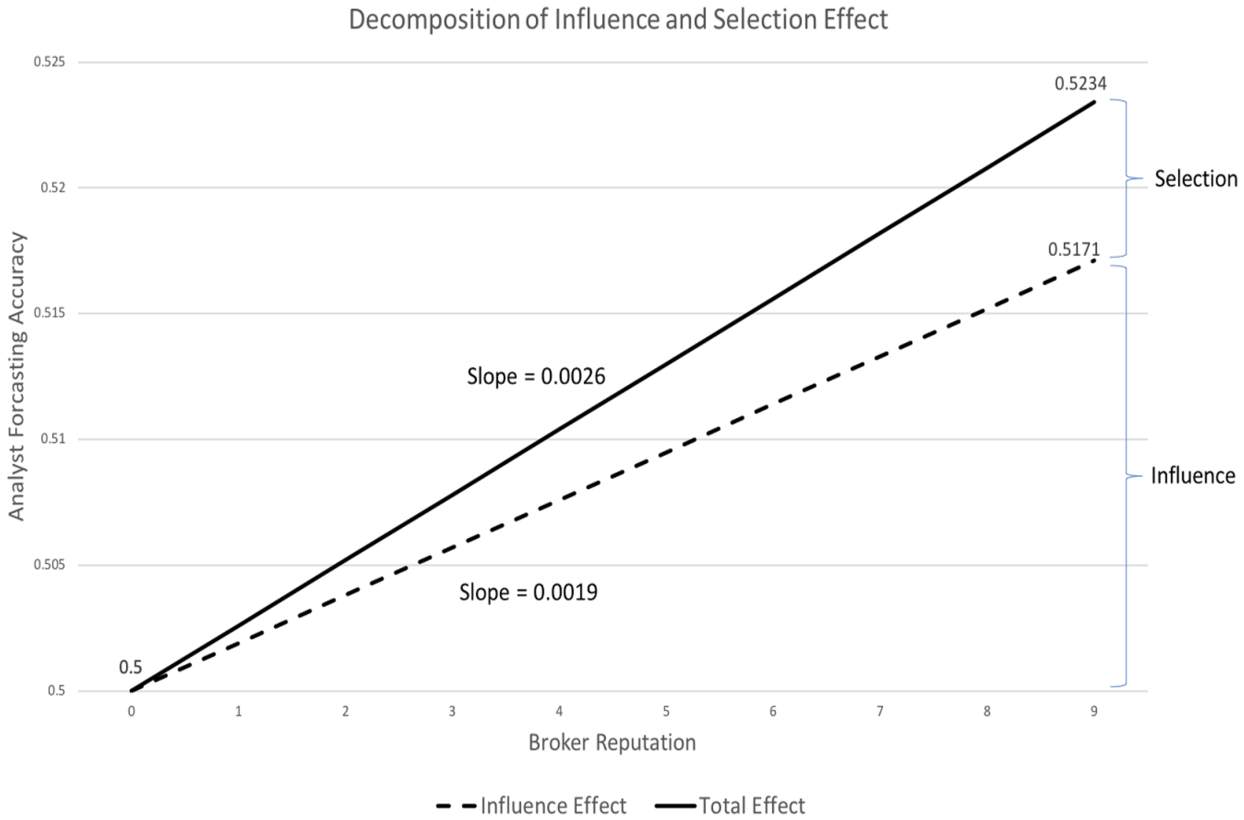


Table I: Summary statistics

This table reports summary statistics of the main variables. We consider an analyst's tenure as her first five years working for the brokerage firm. Broker reputation is the Carter and Manaster rank on a scale of 0 to 9, and the higher the rank the more prestigious the brokerage firm. Broker growth is the percentage of brokerage size increase from last year. Number of stocks and industries is the average number of firms and industries she covers during her tenure. Log(Market Cap) is the logarithm of the total market cap an analyst covers in her first year. Ethnicity indicates whether the analyst is white Caucasian or not based on the analyst's surname (1 indicates not, 0 indicates yes). To include the focus industry fixed effects, we define industries using the Fama-French five industry classifications, and classify an analyst's focus industry as the one in which she covers the most stocks. We indicate the following four industries: Consumer (including retails & wholesales), Manufacturing & Energy, High Tech, and Health. Num IPO indicates the total number of IPOs made in a specific year.

				Percentile				
Variables	N	Mean	St. Dev	10th	25th	50th	75th	90th
Accuracy	1,815	0.514	0.082	0.410	0.473	0.521	0.567	0.609
Broker Reputation	1,815	5.948	3.112	0	5.001	7.001	8.501	9.001
Broker Growth	1,815	0.145	0.467	-0.181	-0.066	0.052	0.191	0.500
Log(Market Cap)	1,815	8.642	1.974	6.052	7.260	8.602	10.119	11.952
Num Stocks	1,815	8.511	4.780	2.6	5	8	11.4	14.75
Num Industries	1,815	1.683	0.771	1	1	1.5	2	2.8
Ethnicity	1,815	0.167	0.374	0	0	0	0	1
I.Consumers	1,815	0.141	0.348	0	0	0	0	1
I.Manuf & Energy	1,815	0.196	0.397	0	0	0	0	1
I.High Tech	1,815	0.279	0.449	0	0	0	1	1
I.Health	1,815	0.116	0.320	0	0	0	0	1
Num IPO	1,815	168.047	148.497	38	60	131	223	384

Table II: Naive OLS Regression

This table reports estimation results of the OLS model for analyst accuracy. Columns (1) to (3) present this relationship by using the whole sample from 1996 to 2013. Column (4) analyzes this relationship using the first half of the sample and column (5) analyzes the relationship using the second half of the sample. Parentheses include the corresponding standard errors. ***,**, and * denote significance at the 10%, 5%, and 1% levels, respectively. Variables are defined in Table I.

	Analyst forecasting accuracy				
	Whole sample			1996 - 2004	2005 - 2013
VARIABLES	(1)	(2)	(3)	(4)	(5)
Broker prestige	0.0025***	0.0025***	0.0026***	0.0022**	0.0031***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Broker growth		0.0005	-0.0032	-0.0045	-0.0017
		(0.005)	(0.005)	(0.006)	(0.008)
Log(Market cap)		0.0006	0.0005	0.0020	-0.0011
		(0.001)	(0.001)	(0.002)	(0.002)
Num stocks		0.0017***	0.0015***	0.0019**	0.0013**
		(0.000)	(0.000)	(0.001)	(0.001)
Num industries		0.0017	0.0022	0.0012	0.0033
		(0.002)	(0.002)	(0.004)	(0.003)
Ethnicity		0.0075	0.0099*	0.0230***	0.0016
		(0.005)	(0.005)	(0.008)	(0.007)
Num IPOs		0.0000**	-0.0002***	-0.0001	-0.0005**
		(0.000)	(0.000)	(0.000)	(0.000)
Market dummy	No	No	Yes	Yes	Yes
Observations	1,815	1,815	1,815	785	1,030
R-squared	0.0088	0.0212	0.0490	0.0570	0.0458

Table III: Bayesian estimate of the matching model and the outcome equation

This table reports Bayesian estimation results of two equations from the structure model. The dependent variable in the outcome equation is analyst forecast accuracy, and the dependent variable in the valuation equation is the latent matching value. A detailed description of the variables is given in Table I. Mean, Median, and Standard Dev. are the statistics of the simulated posterior distributions of the parameters. Marginal effects of the valuation equation represent the probability of choosing two matches with only marginal change in one variable, and are calculated by following Sorensen (2007a). Estimates are based on 110,000 simulations of the posterior distribution. The initial 11,000 simulations are discarded for burn-in. ***,**,and * denote that zeros are not contained in the 10%, 5%, and 1% credible intervals, respectively. Variables are defined in Table I.

	Dependent variable: Analyst forecasting accuracy				
	Mean	Median	Marginal effect	Standard Dev.	95% HPD
VARIABLES	(1)	(2)	(3)	(4)	(5)
Panel A: Outcome equation					
Broker reputation	0.0019***	0.0019		0.0007	[0.0006, 0.0033]
Broker growth	-0.0044	-0.0044		0.0043	[-0.0130, 0.0041]
Log(Market cap)	-0.0002	-0.0002		0.0011	[-0.0024, 0.0019]
Num stocks	0.0013***	0.0013		0.0005	[0.0004, 0.0022]
Num industry	-0.0031	-0.0031		0.0027	[-0.0084, 0.0023]
Ethnicity	0.0036	0.0037		0.0052	[-0.0070, 0.0140]
Num IPO	0.0000*	0.0000		0.0001	[-0.0000, 0.0001]
Panel B: Matching equation					
Broker reputation	0.1439***	0.1409	0.0406	0.0261	[0.0974, 0.1952]
Broker growth	0.0651	0.0663	0.0184	0.1219	[-0.1732, 0.3010]
Log(Market cap)	0.0569***	0.0560	0.0161	0.0149	[0.0284, 0.0868]
Num stocks	-0.0095	-0.0091	-0.0027	0.0068	[-0.0233, 0.0034]
Num industry	-0.1763***	-0.1753	-0.0497	0.0316	[-0.2381, -0.1145]
Ethnicity	0.1820***	0.1786	0.0513	0.0664	[0.0556, 0.3115]
Panel C: Variance					
δ	0.0063***	0.0063		0.0037	[8.89e-07, 0.0131]

Table IV: Bayesian estimate of alternative market and comparison

This table compares the outcome equations from models with different market definitions and compares the coefficient estimated from the naive OLS regression for analyst accuracy. Bayesian estimates are based on 110,000 simulations of the posterior distribution. The initial 11,000 simulations are discarded for burn-in and a tune-in factor of 10. Parentheses represent the corresponding t-statistics. ***,**,and * denote significance at the 10%, 5%, and 1% levels, respectively. Variables are defined as in Table I.

	Dependent variable: Analyst forecasting accuracy				
	OLS	Bayesian estimation		Difference with OLS	
		Main	Expanded market	Main	Expanded market
VARIABLES	(1)	(2)	(3)	(4)	(5)
Broker reputation	0.0026	0.0019	0.0019	0.0007**	0.0007**
				(2.3459)	(2.3459)
Broker growth	-0.0032	-0.0044	-0.0004	0.0012	-0.0028
				(0.5741)	(-1.3397)
Log(Market cap)	0.0005	-0.0002	0.0001	0.0008	0.0005
				(1.4906)	(0.8518)
Num stocks	0.0015	0.0013	0.0009	0.0002	0.0006***
				(0.9377)	(2.8132)
Num industry	0.0022	-0.0031	-0.0034	0.0053***	0.0056***
				(5.3979)	(5.7035)
Ethnicity	0.0099	0.0036	-0.0031	0.0063***	0.0130***
				(2.6854)	(5.5420)
Num IPO	-0.0002	0.0000	0.0000	-0.0002***	-0.0002***
				(-9.5108)	(-9.5108)
δ		0.0063	0.0074		
Markets	36	36	18		

Appendices

A MCMC estimation procedure

Let the markets be indexed by $m = 1, \dots, N$, latent valuation variables be $V_m \equiv \{V_{ij}, ij \in M_m\}$, matching characteristics $W_m \equiv \{W_{ij}, ij \in M_m\}$, and exogenous explanatory variables be $X_m \equiv \{X_{ij}, ij \in M_m\}$, for all potential matches $ij \in M_m$ in each market m. The following algorithm shows how to draw from the posterior the distribution of the parameters augmented with the latent valuation variable, V_{ij}, and the missing observations y^*_{ij} for unobserved matches. We are interested in estimating the parameters α, β, and δ. The Markov chain is generated by drawing each individual dimension of the joint posterior distribution conditional on the draws of the other dimensions as follows:

1. Start Gibbs-sampler for $g = 1 : G_{burn-in} + G_{sample}$ total runs.

2. Initialise the sampling by drawing α, β, δ, and σ^2_ξ from prior distributions: $\alpha \sim N(\alpha_0, A_\alpha^{-1} = 10I_k)$, $\beta \sim N(\beta_0, A_\beta^{-1} = 10I_p)$, $\delta|\sigma^2_\xi \sim N(\delta_0, \sigma^2_\xi/A_\delta)$, and $\sigma^2_\xi \sim IG(a = 2.1, b = 1)$.

3. Draw latent valuation variables V_{ij} for all potential matches in each market m, and draw outcome variable Y_{ij} for unobserved matches in each market m, from distributions conditional on parameters $\alpha, \beta, \delta, \sigma^2_\xi$.

4. Update α, β by drawing from a Bayesian Seemingly Unrelated Regression (BSUR) of $[V; Y]$ on $[W; X]$ conditional on δ, σ^2_ξ.

5. Update δ, σ^2_ξ by drawing from a Bayesian regression of $Y - X\beta$ on $V - W\alpha$, conditional on α, β.

6. Go back to step 3 and repeat.

We now describe how to draw from each conditional distribution.

A.1 Conditional distribution of valuation variables V_{ij}

The conditional augmented posterior distribution of V_{ij} depends on whether brokerage firm i and analyst j are matched or not:

- when $ij \notin \mu_m$, we draw V_{ij} from $N(W'_{ij}\alpha, 1)$ truncated from above at $\overline{V}_{ij}$;

- when $ij \in \mu_m$, we draw V_{ij} from

$$V_{ij}|\alpha, \beta, \delta, \sigma_\xi^2, Y_{ij} \sim N\left(W'_{ij}\alpha + (Y_{ij} - X'_{ij}\beta)\frac{\delta}{\delta^2 + \sigma_\xi^2}, \frac{\sigma_\xi^2}{\delta^2 + \sigma_\xi^2}\right)$$

  truncated from below at $\underline{V}_{ij}$.

The expressions for $\overline{V}_{ij}$ and $\underline{V}_{ij}$ are given in the equation.

A.2 Conditional distribution of unobserved outcome variables Y_{ij}

We only need to simulate the outcome variable Y_{ij} if $ij \notin \mu_m$, i.e., for unobserved matches. We draw Y_{ij} from

$$Y_{ij}|\alpha, \beta, \delta, \sigma_\xi^2, V_{ij} \sim N\left(X'_{ij}\beta + \delta(V_{ij} - W'_{ij}\alpha), \sigma_\xi^2\right).$$

A.3 Conditional distribution of α and β

We apply a BSUR of [V; Y] on [W;X] to sample α and β,

$$\alpha, \beta|V_{ij}, Y_{ij}, \delta, \sigma_\xi^2 \sim N\left(M^{-1}N, M^{-1}\right),$$

where

$$M = \begin{pmatrix} \Omega_{1,1}^{-1}W'W & \Omega_{1,2}^{-1}W'X \\ \Omega_{2,1}^{-1}X'W & \Omega_{2,2}^{-1}X'X \end{pmatrix} + A, \quad N = \begin{pmatrix} A_\alpha\alpha_0 \\ A_\beta\beta_0 \end{pmatrix} + \begin{pmatrix} \Omega_{1,1}^{-1}W'V & \Omega_{1,2}^{-1}W'Y \\ \Omega_{2,1}^{-1}X'V & \Omega_{2,2}^{-1}X'Y \end{pmatrix},$$

and

$$\Omega = \begin{pmatrix} 1 & \delta \\ \delta & \delta^2 + \sigma_\xi^2 \end{pmatrix}.$$

A.4 Conditional distribution of δ and σ_ξ^2

Draw $\delta, \sigma_{xi}^2 | \alpha, \beta, V, Y$ from a Bayesian regression of $\varepsilon = Y - X\beta$ on $\eta = V - W\alpha$:

1. Draw $\sigma_\xi^2 \sim IG(a + N, b + S)$, where N is the number of all potential matches from all markets, and $S = (\varepsilon - \eta d)'(\varepsilon - \eta d) + (d - \delta_0)' A_\delta (d - \delta_0)$, and $d = (\eta'\eta + A_\delta)^{-1}(\eta'\varepsilon + A_\delta \delta_0)$.

2. Draw $\delta | \sigma_\xi^2 \sim N\left(d, \sigma_\xi^2 (\eta'\eta + A_\delta)^{-1}\right)$, truncated from below at 0.

References

Agarwal, N. (2015). An Empirical Model of the Medical Match. *American Economic Review*, 105(7):1939–1978.

Akkus, O., Cookson, J. A., and Hortaçsu, A. (2016a). Endogenous Matching, Underwriter Reputation, and the Underpricing of Initial Public Offerings. *SSRN Electronic Journal*.

Akkus, O., Cookson, J. A., and Hortaçsu, A. (2016b). The Determinants of Bank Mergers: A Revealed Preference Analysis. *Management Science*, 62(8):2241–2258.

Albert, J. H. and Chib, S. (1993). Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association*, 88(422):669–679.

Brown, L. D., Call, A. C., Clement, M. B., and Sharp, N. Y. (2015). Inside the "Black Box" of Sell-Side Financial Analysts. *Journal of Accounting Research*, 53(1):1–47.

Carter, R. and Manaster, S. (1990). Initial Public Offerings and Underwriter Reputation. *The Journal of Finance*, 45(4):1045–1067.

Carter, R. B., Dark, F. H., and Singh, A. K. (1998). Underwriter Reputation, Initial Returns, and the Long-Run Performance of IPO Stocks. *The Journal of Finance*, 53(1):285–311.

Chen, J. (2014). Estimation of the Loan Spread Equation with Endogenous Bank-Firm Matching. In *Structural Econometric Models*, pages 251–289. Emerald Group Publishing Limited.

Chen, J. and Song, K. (2013). Two-sided matching in the loan market. *International Journal of Industrial Organization*, 31(2):145–152.

Chuprinin, O. and Sosyura, D. (2018). Family Descent as a Signal of Managerial Quality: Evidence from Mutual Funds. *Review of Financial Studies*, 27:3133–hhy010 VL – IS –.

Clement, M. B. (1999). Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics*, 27(3):285–303.

Edmans, A. (2011). Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics*, 101(3):621–640.

Fox, J. T., Yang, C., and Hsu, D. H. (2018). Unobserved Heterogeneity in Matching Games. *Journal of Political Economy*, pages 000–000.

Gale, D. and Shapley, L. S. (1962). College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 69(1):9.

Geweke, J. F., Keane, M. P., and Runkle, D. (1994). Alternative Computational Approaches to Inference in the Multinomial Probit Model. *The Review of Economics and Statistics*, 76(4):609–632.

Groysberg, B., Healy, P. M., and Maber, D. A. (2011). What Drives Sell-Side Analyst Compensation at High-Status Investment Banks? *Journal of Accounting Research*, 49(4):969–1000.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161.

Hong, H., Kubik, J. D., and Solomon, A. (2000). Security Analysts' Career Concerns and Herding of Earnings Forecasts. *The RAND Journal of Economics*, 31(1):121–144.

Kothari, S. P., So, E., and Verdi, R. (2016). Analysts' Forecasts and Asset Pricing: A Survey. *Annu. Rev. Financ. Econ.*, 8(1):197–219.

Loughran, T. and Ritter, J. (2004). Why Has IPO Underpricing Changed over Time? *Financial Management*, 33(3):5–37.

Matveyev, E. (2016). The Labor Market for Corporate Directors. *SSRN Electronic Journal*.

Mikhail, M. B., Walther, B. R., and Willis, R. H. (1999). Does Forecast Accuracy Matter to Security Analysts? *The Accounting Review*, 74(2):185–200.

Ni, J. and Srinivasan, K. (2015). Matching in the Sourcing Market: A Structural Analysis of the Upstream Channel. *Marketing Science*, 34(5):722–738.

Pan, Y. (2015). The Determinants and Impact of Executive-Firm Matches. *Management Science*, page mnsc.2015.2278.

Park, M. (2013). Understanding merger incentives and outcomes in the US mutual fund industry. 37(11):4368–4380.

Roth, A. E. and Sotomayor, M. A. O. (1992). *Two-Sided Matching*. A Study in Game-Theoretic Modeling and Analysis. Cambridge University Press.

Schwert, M. (2018). Bank Capital and Lending Relationships. *The Journal of Finance*, 73:1175–n/a.

Soltes, E. (2013). Private Interaction Between Firm Management and Sell-Side Analysts. *Journal of Accounting Research*, 52(1):245–272.

Sorensen, M. (2007a). How Smart Is Smart Money? A Two-Sided Matching Model of Venture Capital. *The Journal of Finance*, 62(6):2725–2762.

Sorensen, M. (2007b). Identification of Multi-Index Sample Selection Models. *SSRN Electronic Journal*.

Stickel, S. E. (1992). Reputation and Performance Among Security Analysts. *The Journal of Finance*, 47(5):1811–1836.

Tanner, M. A. and Wong, W. H. (1987). The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, 82(398):528–540.

Xia, S. (2018). Selection Versus Incentives in Incentive Pay: Evidence from a Matching Model. *SSRN Electronic Journal*.