NBER WORKING PAPER SERIES

AN EMPIRICAL MODEL OF STOCK ANALYSTS' RECOMMENDATIONS:
MARKET FUNDAMENTALS, CONFLICTS OF INTEREST,
AND PEER EFFECTS

Patrick Bajari
John Krainer

An Empirical Model of Stock Analysts' Recommendations: Market Fundamentals, Conflicts of
Interest, and Peer Effects
Patrick Bajari and John Krainer
NBER Working Paper No. 10665
August 2004
JEL No. G3, L1, L5, C1

ABSTRACT

In this paper we develop an empirical model of equity analyst recommendations for firms in the

NASDAQ 100 during 1998-2003. In the model we allow recommendations to depend on publicly

observed information, measures of an analyst's beliefs about a stock's future earnings, investment

banking activity, and peer group effects which determine industry norms. To address the reflection

problem, we propose a new approach to identification and estimation of models with peer effects

suggested by recent work on estimating games. Our empirical results suggest that recommendations

depend most heavily on publicly observable information about the stocks and on industry norms.

In most of our specifications, the existence of an investment banking deal does not have a

statistically significant relationship with analysts' stock recommendations.

Patrick Bajari
219B Social Science Building
Department of Economics
Duke University
Durham, NC 27708-0097
and NBER
bajari@econ.duke.edu

John Krainer
Federal Reserve Bank of San Francisco
101 Market Street
San Francisco, CA 94105
john.krainer@sf.frb.org

1 Introduction

Equity analysts play an important role in stock markets by reducing the duplication of costly effort and by gathering information necessary for the efficient pricing of stocks. For as long as there have been equity analysts, however, there have been questions about the accuracy of their forecasts and their implicit buy and sell recommendations (see Cowles (1933)). This is particularly true in the wake of the sharp stock price declines for technology firms in 2000. Not only did recommended stocks vastly underperform the market as a whole during this period, but highly-publicized allegations of conflicts of interest have called into question whether analysts were more concerned with helping their firms win investment banking business than with producing accurate assessments of the prospects for the firms under scrutiny.

In this paper, we develop an empirical model of the recommendations generated by stock analysts. We quantify the relative importance of four factors influencing the production of recommendations in a sample of high technology stocks during the time period between 1998 and 2003. First, recommendations must depend on fundamentals and commonly shared expectations about the future profitability of the firm. These expectations will be embedded in the stock price. Second, analysts are heterogeneous, both in terms of talent and perhaps in terms of access to information. We try to capture an individual analyst's private belief about the stock by looking at the difference between the quarterly earnings forecast submitted by the analyst (or the analyst's brokerage firm) and the distribution of forecasts from other firms. Mindful of the large number of inquiries into possible conflicts of interest among research analysts, we include as a third

2

factor a dummy variable for an investment banking relationship between the firm and the analyst's employer. Finally, we consider the influence of peers on the recommendation decision. Peer effects can impact the recommendation in different ways. Individual analysts have incentive to condition their recommendation on the recommendations of their peers, because even if their recommendations turn out to be unprofitable ex-post, performance evaluation is typically a comparison against the performance of peers. More subtly, recommendations are relative rankings of firms and are not easily quantifiable (or verifiable) objects. As such, ratings scales usually reflect conventions and norms. The phenomenon is similar to the college professor's problem of assigning grades. If a professor were to award the average student with a C while other faculty give a B+ to the average student, the professor might incorrectly signal his views of student performance. Even while there is heterogeneity in how individual professors feel about grading, most conform to norms if only to communicate clearly with students (and their potential employers) about their performance. Similarly, analysts have an incentive to benchmark their recommendations against perceived industry norms.

The choice of a recommendation is naturally modeled as a discrete choice problem. It is important to recognize, however, that if analysts benchmark their recommendation to the recommendations of their peers, then the choice of one analyst is simultaneously determined with the choices of other analysts. Therefore, our principal econometric model is a system of simultaneous discrete choice models. Recently, industrial organization economists have proposed computationally simple estimators that can be applied to discrete games. Examples include Bresnahan and Reiss (1990, 1991), Berry (1992), Tamer (2002), Seim (2003), Aguirregabiria and Mira (2003), Berry and Pakes (2002), Pakes, Ovstrovsky, and Berry (2003), Bajari, Benkard and Levin (2003), Pesendorfer and Schmidt-Dengler (2003), Bajari, Hong and Ryan (2004) and

3

Sweeting (2004). Following the literature, we propose a simple two-step estimator for this game.

This paper makes three contributions to the literature. First, our paper is one of the first applications of new techniques for estimating discrete games (see Aguirregabiria and Mira (2003), Pakes, Ovstrovsky, and Berry (2003), Bajari, Benkard and Levin (2003) and Pesendorfer and Schmidt-Dengler (2003)). These methods allow for a flexible specification of the game but are simple to compute.

Second, we propose methods to identify the parameters of our model under weak functional form assumptions. The identification strategy is similar to Bajari, Hong and Ryan (2004), but applied to a different class of models. We demonstrate that two types of exclusion restrictions are sufficient for identification. First, we search for variables that enter into the utility of a particular player, but which can be excluded from the utility of other players. Shifts in these variables change the utility of a single player holding the rest of the model parameters fixed. Second, the models that we consider can generate multiple equilibria. We demonstrate the model is identified if there are some variables that influence the selection of equilibria, but otherwise can be excluded from the utility. We argue that the threat of sanctions by regulators such as New York State Attorney General Elliot Spitzer served as instruments that encouraged analysts to coordinate on an equilibrum where more conservative recommendations were issued.

The third contribution is that, to the best our knowledge, we are the first empirical paper to study biases on recommendations from conflicts of interest for high technology stocks in the late 1990's. The perceived conflicts of interest were widely reported on during this time. Public confidence in the markets was damaged by the perception that analysts encouraged investors to buy stocks that they were privately disparaging. However, the empirical literature has not quantified the bias from conflicts of interest during this episode.

There is now a huge finance literature on the analyst's role in the information production process and

the question of whether or not analyst recommendations are valuable. Our paper touches on this literature, even though our methods are quite different. In theory, we would expect analyst recommendations to have value, otherwise they would not be produced (see Grossman and Stiglitz (1980)). However, the empirical literature has generated a mixed picture about the direct value of analyst information. Cowles (1933) was one of the first researchers to question whether investors could trade profitably on analyst recommendations, and concluded that they could not. In general, it was found that abnormal returns around a change in recommendation were either very small, or mean reverting.[2]

In the last ten years, however, the finance profession has revisited this question and found evidence of investment value associated with the analysts. Using data in the 1989-1991 time period, Womack (1996) reports mean price increases of approximately 2.4% for buy recommendations and -9.1% for sell recommendations. Barber, Lehavey, McNichols, and Trueman (2001) come to a similar conclusion using data from the 1989-1996 period and a slightly different empirical strategy. Barber et. al. (hereafter BLMT) find that a strategy of purchasing the most highly recommended stocks outperformed holding the least recommended stocks by almost 1% per month. This finding turned out to be robust after controlling for transaction costs, risk, and other attributes in the recommended portfolios.

The paper is organized as follows. In section 2 we describe the environment for Internet stocks in the late 1990s and how it was difficult for brokerage firms to avoid perceptions that their dealings were tainted by conflicts of interest. In section 3 we describe the data. In section 4 we outline the model. In section 5 we study the identification of the model. In section 6 we report the results. Section 7 concludes.

---

[2] The discouraging performance of the analysts in the *Wall Street Journal's* popular "dartboard" competition is a persistent reminder of this point.

2 The tech boom and analyst conflicts of interest

The latter 1990's were witness to a dramatic rise and then subsequent decline in stock prices. This volatility was particularly pronounced in the technology-dominated NASDAQ. In January 1999, the NASDAQ composite index was slightly over 2000. The index reached a high of over 5,000 about a year later, and then fell to 1,300 by November of 2002. The firms listed on the NASDAQ are typically young firms that are engaged in innovative but unproven lines of business. Relatively little is known about the demand for the firms' products and future performance is hard to predict from past performance. This would seem to have been an ideal environment for equity analysts to enter, provide information, and add investment value. It is puzzling that this did not happen at this key point in time. With hindsight, it appears obvious now that stocks were priced with unrealistic expectations for earnings growth. Yet, the average equity analyst appears to have been as optimistic as any market participant leading up to the crash in the NASDAQ. BLMT (2003) updated their earlier study on the investment value of analyst in the early 1990s to show that analyst recommendations were much less valuable during 1996-2001 time period. In 2000 and 2001, the least recommended stocks earned an average abnormal return of 13%, while the most highly recommended stocks earned average abnormal returns of -7%.

BLMT note that the list of most highly recommended stocks in 2000-2001 contained an inordinate number of small cap stocks, which turned out to be the ones that fell most in value. These firms, of course, are usually young and have relatively more demand for investment banking services than more mature firms. Some researchers have established the link between certain types of ratings outcomes and investment banking.[3] Lin and McNichols (1998) document that analysts tend to give more favorable ratings

---

[3] BLMT do not actually speak to whether conflicts of interest could be responsible for their results.

and more optimistic long term growth forecasts to firms when their own company serves as an underwriter to the firm. Michaely and Womack (1999) show how stocks recommended by a brokerage firm that also managed the initial public offering of the stock lead to significant underperformance relative to unaffiliated analyst recommendations.[4] This explanation of analyst behavior gained even more credence when several securities firms were forced to disclose internal documents that seemed to show investment banker influence on the analyst recommendation. Elliot Spitzer, the New York State Attorney General, issued a complaint on the matter which eventually led to a $600 million dollar settlement involving a host of the leading Wall Street investment banks.

Unlike many papers in this literature, we choose to focus our analysis on a relatively short time period and a small sample of firms. We focus on a period (the late 1990s) and a set of stocks (technology firms) where the role of the analyst was alleged to have been most dysfunctional. We do this for several reasons. First, we believe that the influence of norms and peers is likely to be greatest in cases where there is substantial uncertainty about fundamentals. Second, if we are to differentiate between conflicts of interest and other explanations, we require a time period with variation in the pressures from these alleged conflicts. Once Elliot Spitzer's investigation was public knowledge, there was surely less pressure on analysts to boost their recommendations. It turns out that the timing of the Spitzer investigation serves nicely as an instrument for estimating our model.

3 The data

The data consist of the recommendations on firms that made up the NASDAQ 100 index as of year-end

---

[4] Other papers that examine conflicts of interest between investment banking and research analysis include Ali (1996), Chan, Karceski, and Lakonishok (2003), Dugar and Nathan (1995), Francis and Philbrick (1993), and Korajczk, Lucas and McDonald (1991).

2001. The recommendations were collected from Thomson Firstcall. Firstcall is one of the most comprehensive historical data sources for analysts' recommendations and earnings forecasts, containing real-time recommendations and forecasts from hundreds of analysts. It is common for analysts to rate firms on a 5 point scale, with 1 denoting the best recommendation and 5 denoting the worst. When this is not the case, Firstcall converts the recommendations to the 5 point scale (see table 3.1).

We have 12,719 recommendations from analysts at 185 brokerage firms between January 1998 and June 2003 (see Table 2). The data set was formed by merging the earnings and recommendations files from Firstcall. In a given quarter, for a given stock, we merge a quarterly earnings forecast with a recommendation from the same brokerage.[5] This will allow us to determine if analysts that are more optimistic than the consensus tend to give higher recommendations. In the Firstcall data, quarterly earnings forecasts are frequently made more than a year in advance. In order to have a consistent time frame, we limit analysis to forecasts that were made within the quarter that the forecast applies.[6] Note that not every recommendation can be paired with an earnings forecast made within the contemporaneous quarter. Recommendations that could not be paired with an earnings forecast were dropped from the results that we report. However, qualitatively similar results were found for a data set where this censoring was not performed. We choose not to report these results in the interests of brevity. The variables in our data include:

- REC- Recommendation from 1-5 for a stock listed in the NASDAQ 100 recorded by I/B/E/S.
- QUARTER- Quarter during which the recommendation was submitted.
- STOCK-Name of the stock for which the recommendation applies.
- BROKERAGE-The brokerage employing the analyst.
- EPS-Earnings per share forecast submitted by the analyst's brokerage associated with the recommendation. Submitted during the same quarter as the recommendation.

---

[5] When there were multiple recommendations within a quarter, we chose to use the last recommendation in the results that we report.
[6] We chose to merge the brokerage field, instead of the analysts field, because the names and codes in the analysts field were not recorded consistently across IBES data sets for recommendations. It was possible to merge at the level of the brokerage.

- AEPS-Average of the earnings per share forecasts submitted for that quarter.

- RELATION-A dummy variable that is one if the analyst's brokerage engages in investment banking business with the company to which the recommendation applies.

- IBANK-A dummy variable that is equal to one if the brokerage does any investment banking business with stocks in the NASDAQ 100.

- SPITDUM-A dummy variable that is equal to one after the quarter starting in June of 2001. Based on a comprehensive search of Wall Street Journal articles, this is when Elliot Spitzer began making very public criticisms of industry practices.

- SBANK-the share of analysts that issued recommendations for a particular stock during a particular quarter where IBANK was one.

The investment banking relationship was identified from several different sources. First, we checked form 424 filings in the SEC's database for information on the lead underwriters and syndicate members of debt issues. When available, we used SEC form S-1 for information on financial advisors in mergers. We also gathered information on underwriters of seasoned equity issues from Securities Data Corporation's Platinum database. Transaction advisory services (mergers), and debt and equity issuance are not the only services that investment banks provide. However, these sources contribute the most to total profitability of the investment banking side of a brokerage firm.

We provide some summary statistics for the variables in our data set in Table 3.2. The average recommendation in our data set is 2.2, which is approximately a buy recommendation. The mean value of RELATION is 0.035. The mean valued of IBANK is 0.81. That is, three and one-half percent of the analysts-company pairs in our data set were identified as having a potential conflict of interest due to investment banking activity. Eighty-one percent of the recommendations in our data were generated by firms engaging in investment banking with at least one firm listed in the NASDAQ 100. Both of these variables are potentially useful measures of potential conflict of interest. The variable RELATION is more direct, since it indicates that the brokerage is engaged in investment banking with the company during the quarter

the recommendation was issued. However, brokerages might view any company it is giving a recommendation to as a potential client, particularly in the NASDAQ 100, where many of the companies generated considerable investment banking fees.

3.1 Factors influencing recommendations

In this subsection, we discuss variables that could influence analysts' recommendations. These variables will enter into our empirical model in section 6.

The first hypothesis we entertain is that recommendations are influenced by the analyst's beliefs about the fundamental worth of the company. In our data set, three variables will proxy for the analysts beliefs about the company fundamentals. The first variable is the difference between the analyst's earnings forecast and the average earnings forecast issued within that quarter. If an individual analyst's earnings forecast is greater than the average, we interpret this to mean that he is more optimistic about the company's future earnings than are his peers. If, all else held equal, analysts recommend stocks with higher earnings prospects, then we would expect stocks that an analysts' brokerage is more optimistic about to be more highly recommended. The second and third variables are stock and time fixed effects, which presumably capture publicly available information about the stocks which could influence recommendations. Ideally, we would like to use valuation measures such as book-to-market and the P/E ratio. However, for the NASDAQ 100 during the height of the technology stock boom, expectations about the growth of the market and profitability in possibly distant quarters may have been more important in determining both market prices and recommendations. Given the breakdown in the relationship between stock prices and more standard fundamentals, capturing publicly observed information about stocks through the use of stock and time fixed effects seems to us a more appropriate empirical strategy.

A second factor that could influence recommendations is the potential conflict of interest posed by the allure of capturing investment banking fees. Media reports and revelations from Elliot Spitzer's investigation lead one to suspect that some individual stock analysts may have played a key role in helping their firms win investment banking business during this period.[7] Yet, conflicts of interest were not something new to this period. Indeed, Michaely and Womack (1999) report that in 1990-91 a buy recommendation by an analyst whose firm took the stock public was significantly less valuable than a buy recommendation by an unaffiliated analyst. What is interesting about our data set is that the conflict of interest temptation was particularly acute in the early part of our sample because of the large demand for corporate finance at this time for this set of firms. In effect, variation in the conflict of interest variable gives us a better chance of precisely estimating its effect and the effects of the other variables in the model.

One key issue is how to construct the investment banking variable. Firms change their investment bankers from deal to deal. Having once had an investment banking relationship does not necessarily imply that an analyst must forever bias his recommendation. Accordingly, we measure the investment banking relationship with a dummy variable equal to one if we can establish an investment banking relationship either one year prior to or one year after a recommendation. We did not find the results to be sensitive to this assumption.

The third factor which could influence the choice of recommendation is a peer effect. Table 3.3 illustrates that the distribution of recommendations moved considerably over our sample period. While part of this movement can be traced to covariates, we will argue that some of it is due to the inherent ambiguity of

---

[7] In 1998, Goldman Sachs estimated that Jack Grubman, a prominent telecommunications industry analyst, would bring in $100 to $150 million in investment banking fees. This estimate was based on the fees generated by 32 of the stocks he covered that also had banking relationships with Citigroup, including WorldCom, Global Crossing and Winstar Communications. (Wall Street Journal, October 11, 2002).

assigning grades.

As an example, consider the problem of assigning grades in the classroom. There is considerable heterogeneity across schools in how grades are submitted. For instance, in United States graduate programs in economics, it is rare to give any student a grade of "C" on coursework, particularly after the first year. However, in Law Schools, by comparison, grades of "C" and "D" are fairly common. Just like grades differ substantially across schools and over time, different analysts (or brokerage firms) will have different approaches to grading companies. These different approaches may also change over time in ways not related to the fundamentals driving the performance of firms under analysis.

The key to our modeling approach is the observation that analyst recommendations are not wholly independent of each other even after taking into account their publicly-available information. That is, recommendations should be modeled as a game. If one analyst is uniformly harsher in his evaluations than are his peers, then he risks sending an incorrect signal about the firm under analysis. The heterogeneity in the distribution of recommendations suggests that this game might have multiple equilibria. In the latter part of our data set, well after the precipitous decline in NASDAQ stock prices, the distribution of recommendations shifted so that nearly 15 percent of the stocks received a recommendation of 4 or 5. Almost no stocks had such poor recommendations in the early part of the sample, or even at the time of the steep decline in prices.

4 Identification and Estimation of Static Models of Strategic Interaction

In this section, we describe a strategy for the identification and estimation of a static model of strategic interaction. The estimation strategy is a straightforward application of recent work on estimating games by Aguirregabiria and Mira (2002), Pakes, Ovstrovsky and Berry (2003), Bajari, Benkard and Levin (2003),

and Pesendorfer and Schmidt-Dengler (2003). The specific estimator that we propose will be a parametric implementation of these ideas. The parametric assumptions are used primarily because the number of covariates in our data make nonparametric methods inappropriate. However, it is useful to study whether the identification of the model depends on parametric restrictions. Therefore, we consider nonparametric identification of the model. The identification results are closely related to Bajari, Hong, and Ryan (2004). We include them in this document for completeness and to demonstrate how the strategies in Bajari, Hong and Ryan (2004) can be extended to games of incomplete information.[8]

The model is a simultaneous move game of incomplete information. There are $i = 1, ..., N$ players, each with a finite set of actions A_i. Define $A = \prod_i A_i$ and let $a = (a_1, ..., a_N)$ denote a generic element of A. Player i's utility takes the following form:

$$u_i(a_i, a_{-i}, x, \varepsilon) = f_i(x, a) + \varepsilon_i(a_i). \tag{1}$$

In equation (1), utility depends on two terms. The first is $f_i(x, a)$, a deterministic function that depends on the covariates x and the actions a. The second term, $\varepsilon_i(a_i)$ is a random preference shock that influences i's utility from the action a_i. We assume that the $\varepsilon_i(a_i)$ are draws from a distribution G that are i.i.d. across actions and players. We assume that the $\varepsilon_i(a_i)$ are private information to each player.

Notice that (1) generalizes standard discrete choice models such as the multinomial logit or probit. In these models, it is assumed that f_i does not depend on a_{-i}, the actions of other agents. A standard discrete choice model is a special case of our framework where the number of players in the game is equal to one.

---

[8] Bajari, Hong and Ryan (2004) consider games of complete information with multiple equilibrium. While the strategy of the proofs is similar, the model we consider here has incomplete information and different implicit assumptions about equilibrium selection. At a minimum, we believe that these arguments illuminate the ideas in Bajari, Hong and Ryan (2004) in a simpler context.

4.1 Identifying Assumptions from Discrete Choice

It is well known that (1) cannot be identified even in the special case of a standard probit model. Because choices depend on the differences in the level of utility, there are standard identifying assumptions from discrete choice that we will impose. The first identifying assumption we make is:

A1. For every i and $a_{-i} \in A_{-i}$, we let $f_i(\underline{a_i}, a_{-i}, x) = 0$ for some chosen $\underline{a_i} \in A_i$ and for all $a_{-i} \in A_{-i}$.

The rationale for A1 is similar to the argument that we can normalize the mean utility from the outside good equal to zero in a standard discrete choice model. In order to show that this assumption is without loss of generality, we must verify that it does not change i's ranking of over elements of A. Fix two elements $a^* = (a_1^*, ..., a_N^*)$ and $a' = (a_1', ..., a_N')$ in A. Suppose that $u_i(a^*) \geq u_i(a')$. Then

$$f_i(a^*, x) + \varepsilon_i(a_i) \geq f_i\left(a_i, a_{-i}^*, x\right) + \varepsilon_i(a_i)$$

This inequality will not be affected by subtracting $f_i(\underline{a_i}, a_{-i}, x)$ from both sides for all i. That is:

$$f_i(a^*, x) - f\left(\underline{a_i}, a_{-i}^*, x\right) + \varepsilon_i(a) \geq f_i\left(a_i, a_{-i}^*, x\right) - f\left(\underline{a_i}, a_{-i}^*, x\right) + \varepsilon_i(a_i)$$

Hence assumption A1 does not change the order on A.

A second assumption that we will make is:

A2. For every i and for every a, $\varepsilon_i(a)$ are distributed i.i.d. extreme value.

In standard discrete choice models, it is not possible to identify both $f_i(a, x)$ and $\varepsilon_i(a)$ nonparametrically. We note that distributional assumptions on the error term are necessary even in a binary choice model. Consider a standard binary choice model where the dependent variable is 1 if the index $u(x) + \varepsilon$ is

greater than zero, i.e.

$$y = 1(u(x) + \varepsilon > 0) \tag{2}$$

Suppose that the economist observes $P(y = 1|x)$, the probability that the dependent variable is equal to one given the covariates x. If the cdf of ε is G, then (2) implies that:

$$P(y = 1|x) = G(u(x)), \tag{3}$$

Equation (3) implies that it is not possible to separately identify G from u. In a binary choice model, we can perfectly rationalize by assuming that $u(x) = G^{-1}\left(P(y = 1|x)\right)$. Thus, in the binary case, this assumption will be required for identification. To see a discussion of the multivariate case, see Bajari, Hong and Ryan (2004).

We could change A2 to another standard parametric distribution used in the literature. Similar results would hold in ordered models (such as the ordered logit or ordered probit) or models with different error structures, such as the multinomial probit. The logit model is useful because choice probabilities can be expressed in a closed form. This greatly simplifies the notation and hence allows the reader to understand the key ideas in our approach to identification and estimation.

4.2 Equilibrium

Let $P_i(a_i|x)$ denote the probability that player i takes the action a_i conditional on the covariates x. Then the expected utility that player i receives from taking the action a_i is equal to:

$$\sum_{a_{-i}} f_i(a_i, a_{-i}, x) P_{-i}(a_{-i}|x) + \varepsilon_i(a_i) \tag{4}$$

$$\text{where } P_{-i}(a_{-i}|x) = \prod_{j \neq i} P_j(a_j|s). \tag{5}$$

In equation (4) player i computes his utility from action a_i by marginalizing out over his expectations about the actions of other players, a_{-i}. We note that because assumption A2, conditional on values of $\varepsilon_i(a_i)$, $a_i \in A_i$ there is a unique best response for agent i with probability one. Therefore, the following system of equations would have to be satisfied in a Bayes-Nash equilibrium.

$$P_i(a_i|x) = \frac{\exp(\sum_{a_{-i}} f_i(a_i, a_{-i}, x) P_{-i}(a_{-i}|x))}{\sum_{a_i'} \exp(\sum_{a_{-i}} f_i(a_i, a_{-i}, x) P_{-i}(a_{-i}|x))}$$

We therefore define equilibrium as follows.

Definition. A Bayes-Nash equilibrium is a collection of probability distributions, $P_i^*(a_i|x)$ such that for all i and x

$$P_i^*(a_i|x) = \frac{\exp(\sum_{a_{-i}} f_i(a_i, a_{-i}, x) P_{-i}^*(a_{-i}|x))}{\sum_{a_i'} \exp(\sum_{a_{-i}} f_i(a_i, a_{-i}, x) P_{-i}^*(a_{-i}|x))} \tag{6}$$

$$\text{where } P_{-i}^*(a_{-i}|x) = \prod_{j \neq i} P_j^*(a_j|x) \tag{7}$$

The model of equilibrium in (6)-(7) has received considerable attention in the literature. See, for instance, McKelvy and Palfrey (1995), Seim (2002), Brock and Durlauf (2001), Bayer, McMillan and Rueben (2003), Aguirregabiria and Mira (2003), Pakes, Ovstrovsky and Berry (2003), Pesendorfer and Schmidt-Dengler (2003) and Sweeting (2004). These papers establish the existence of equilibrium using standard fixed point arguments. However, equilibrium is typically not unique. Multiplicity is common in these games and has been studied in detail by McKelvy and Palfrey (1995), Brock and Durlauf (2001,2003) and Sweeting (2004).

4.3 Identification

The problem of identification is to recover the mean utilities, $f_i(a_i, a_{-i}, x)$ given the distributions $P_i(a_i|x)$

16

that in principal can be observed by the economist. We begin by noting that, conditional on x, we can rewrite (6) as follows:

$$\text{For all } i \text{ and any } a_i', a_i \in A_i, \ \log\left(P_i(a_i'|x)\right) - \log\left(P_i(a_i|x)\right) \tag{8}$$

$$= \sum_{a_{-i}} f_i(a_i', a_{-i}, x) P_{-i}(a_{-i}|x)) - \sum_{a_{-i}} f_i(a_i, a_{-i}, x) P_{-i}(a_{-i}|x))$$

In equation (8), we simply transform the choice probabilities implied by the logit model. This equation tells us that conditional on x, we can write the mean utilities as the solution to a linear system in the choice probabilities and the log choice probabilities. For a given x, there are only $N1 = \sum_i \#A_i - N$ such equations that are not linearly independent. For each player, there are $\#A_i$ choice probabilities, however, these probabilities must sum to one. It is this identity which introduces a linear dependence. The number of values needed to characterize $f_i(a_i, a_{-i}, x)$ is equal to $N2 = \sum_i (\#A_i - 1) \prod_{j \neq i} \#A_j$ after we make the normalization implied by A1. By collecting terms, we can therefore express (8) as:

$$A(x) = B(x)f(x) \tag{9}$$

In equation (9), the matrix $A(x)$ is $N1 \times 1$ and is formed by collecting the terms $\log\left(P_i(a_i'|x)\right) - \log\left(P_i(a_i|x)\right)$. The matrix $B(x)$ is $N1 \times N2$ and is formed with terms of the form $P_{-i}(a_{-i}|x))$. The mean utilities are collected in the matrix $f(x)$ which is $N2 \times 1$.

Definition. The model is identified if the $f_i(a, x)$ that satisfy (8)-(9) are uniquely determined by $P_i(a_i|x)$.

Identification means that the probabilities that we observe in the data, $P_i(a_i|x)$ are sufficient to determine $f_i(a, x)$. There is an implicit assumption about the selection of equilibrium in our definition of

17

identification. Given a value of x, there is at most one equilibrium played in the data. In principal there could be multiple equilibria and agents could be randomly switching back and forth between the various equilibria. This would complicate the identification arguments and the estimation. We study this type of framework in Bajari, Hong and Ryan (2004). However, in this paper, we wish to include a very large number of covariates, which would make the approach of Bajari, Hong and Ryan (2004) computationally difficult. Therefore, we consider this more stylized model.[9]

The first result that we establish is that without restrictions on payoffs, the model is not identified.

Theorem 1 *Suppose that A1 and A2 are satisfied. If $N > 1$ and if for all i, $\#A_i > 1$ then $f_i(a, x)$ is not identified.*

Proof: Holding x fixed, the number of moments are equal to $\sum_i \#A_i - N$ which is less than the number of parameters $\sum_i (\#A_i - 1) \prod_{j \neq i} \#A_j$. Q.E.D.

This identification result is similar to many nonidentification results in the literature. See for instance, Bresnahan and Reiss (1990,1991), Pesendorfer and Schmidt-Dengler (2003) and Bajari, Hong and Ryan (2004). The basic problem is that the number of mean utility parameters, conditional on x, far exceeds the number of moments available to the economist. Therefore, the model is not identified.

In order to identify the model, we must either increase the number of equations or decrease the number of unknowns. Our solution to this problem is to impose exclusion restrictions on the $f_i(a, x)$. The first type of exclusion restriction that we use is below.

    A3. For each agent i, there exists some continuous covariate, s_i, that enters i's utility, but not the utility of other agents. That is, i's utility can be written as $f_i(a, x, s_i)$. We let $s = (s_1, ..., s_N)$ denote the vector of shifters for all agents.

---

[9]   We note, however, that many of the identification arguments that we suggest would have analogues in the framework of Bajari, Hong and Ryan (2004).

Assumption A3 implies that there are agent i specific utility shifters. In our investment banking example, this could include the amount of investment banking business done by the firm or some other brokerage-specific covariates. Our assumption implies, for instance, that investment banking work done by Merrill Lynch does not directly influence the utility of Goldman Sachs for issuing a particular recommendation. To a first approximation, we believe this is a reasonable assumption.

In what follows, it will be useful to make the following assumption.

A4. Consider any finite number M of distinct values of (x, s), e.g. $(x^{(1)}, s^{(1)}), ..., (x^{(M)}, s^{(M)})$. Consider the matrix $B\left((x^{(1)}, s^{(1)}), ..., (x^{(M)}, s^{(M)})\right) = \begin{pmatrix} B(x^{(1)}, s^{(1)}) \\ \vdots \\ B(x^{(M)}, s^{(M)}) \end{pmatrix}$ defined analogously to (9).
The matrix $B\left((x^{(1)}, s^{(1)}), ..., (x^{(M)}, s^{(M)})\right)$ has full rank.

Theorem 2 *Suppose that A1-A4 hold. Then f_i is identified.*

Proof: Hold x fixed. Consider a large, but finite number of values of s_i equal to M for each agent. Consider the all the distinct vectors of the form $(x, s_1, ..., s_N)$ that can be formed by allowing for all possible permutations of the individual s_i. The number will be equal to K^N. Consider the moments generated by equation (9) generated by these K^N distinct covariates. The number of moments is equal to $N1 \cdot K^N$. The number of mean utility parameters is equal to $\sum_i K\left(\#A_i - 1\right) \prod_{j \neq i} \#A$. Note that the number of moments depends linearly on K but the but the number of moments grows exponentially with K. By choosing sufficiently large values for K the mean von Neumann-Morgenstern utilities can be identified at $(x, s_1, ..., s_N)$. Q.E.D.

The intuition behind the theorem above is quite simple. Because we have zeroed out the some elements of f_i, by considering all permutations of the $(s_1, ..., s_N)$ we can increase the number of moments at a

squared rate and the number of parameters at a linear rate. It is easy to see that the conditions of the

theorem could be weakened slightly. All we need is that there are individual specific shifters for two of the

players in the game. This is sufficient to increase the number of moments to allow for identification of the

mean utilities. This is similar to an identification strategy in Bajari, Hong and Ryan (2004), but we apply it

here to games of incomplete information.

A second approach to identification is to exploit the multiplicity of equilibria. Previous authors, such

as McKelvy and Palfrey (1995) and Brock and Durlauf (2001,2003), have demonstrated that models of the

form that we consider can generate multiple equilibria.

> **A5.** There is a value x_0 such that for an open set of values $O(x_0)$, the model generates multiple equi-librium. Also, suppose that the probability distributions can be written as $P_i(a_i|x, z)$ where z is a continuous variable which influences which equilibrium is played, but which does not enter payoffs.

As we mentioned earlier in the paper, these variables could include behavior in previous plays of the game

or actions by market regulators. We discuss instruments appropriate for our application in the next section.

If there is such a variable, then we can write (6) as:

$$P_i^*(a_i|x, z) = \frac{\exp(\sum_{a_{-i}} f_i(a_i, a_{-i}, x)P_{-i}^*(a_{-i}|x, z))}{\sum_{a_i'} \exp(\sum_{a_{-i}} f_i(a_i, a_{-i}, x)P_{-i}^*(a_{-i}|x, z))}$$

This implies that we can rewrite (8)-(9) in this case as:

$$\log\left(P_i(a_i'|x, z)\right) - \log\left(P_i(a_i|x, z)\right) = \sum_{a_{-i}} f_i(a_i', a_{-i}, x)P_{-i}(a_{-i}|x, z) - \sum_{a_{-i}} f_i(a_i, a_{-i}, x)P_{-i}(a_{-i}|x, z) \quad (10)$$

$$A(x, z) = B(x, z)f(x) \quad (11)$$

It will also be convenient to make the following assumption.

20

A6. Consider any finite number M of distinct values of $(x, z), (x_1, z_1), ..., (x_M, z_M)$ where $x_i \in O(x_0)$, the set of covariates where there are multiple equilibrium. Consider the matrix $B((x_1, z_1), ..., (x_M, z_M)) = \begin{pmatrix} B(x_1, z_1) \\ \vdots \\ B(x_M, z_M) \end{pmatrix}$. The matrix $B(x_1, ..., x_M)$ has full rank.

Theorem 3 *Suppose that A1-A3 and A5-A6 are true. Then f_i is identified in the set $O(x_0)$.*

Proof: Hold x fixed and consider $\#z$ distinct values of z. Then the number of moments generated by the model is equal to $N1 \cdot \#z$. The number of parameters is equal to $N2$. Since the number of parameters is independent of z, if we let $\#z$ become sufficiently large, the number of moments exceeds the number of parameters and the f_i are identified. Q.E.D.

The intuition behind the identification is similar to Sweeting (2004). The multiplicity of equilibrium generates additional moments that can assist with identification. However, unlike Sweeting, in order to ensure identification in our framework, it is necessary to have an exclusion restriction. Sweeting (2004) does not need this condition because of he assumes a dichotomous choice and makes parametric assumptions about the form of the mean utilities.[10]

4.4 Estimation

Methods for estimating models such as the one considered above have recently been explored by a number of authors including Aguirregabiria and Mira (2003), Pakes, Ovstrovsky, and Berry (2003), Bajari, Benkard and Levin (2003), and Pesendorfer and Schmidt-Dengler (2003). The basic strategy is to construct a two-step estimator. The economist first flexibly estimates $P_i(a_i|x, s, z)$, the probability distribution of the observed actions conditional on the covariates x, the vector of individual shifters s and the variables

---

[10] Bajari, Hong and Ryan (2004) also study identification when there is a variable that influences equilibrium selection, but does not directly enter into utility.

associated with an equilibrium shift z. We estimate $f_i(a, x)$ in a second step.

In principal, nonparametric estimation of such a model is possible. However, in our application, the number of covariates is sufficiently large so that nonparametric methods are not appropriate, due, for instance, to the inclusion of quarter, stock and broker fixed effects that we will discuss in the next section. Therefore, instead, we will use parametric methods.

Let $t = 1, ..., T$ denote an observation in our data set. Let a_t, x_t, s_t and z_t denote the vector of actions and the covariates for observation t. The approach we will take is to first, flexibly estimate a parametric model $P_i(a_i|x, s, z, \theta_1)$ of the probability that player i takes action a_i. This will be done, in our application, by a linear regression. In the second step, given estimates of the vector of parameters, $\widehat{\theta}_1$, and fitted values of the choice probabilities, $\widehat{P}_i(a_i|x, s, z, \widehat{\theta}_1)$, we will maximize the following likelihood in θ_2:

$$L(\widehat{\theta}_1, \theta_2) = \sum_{t=1}^{T} \sum_{i=1}^{N} \log \left(\frac{\exp(\sum_{a_{-i}} f_i(a_{i,t}, a_{-i,t}, s_{i,t}, x_t; \theta_2) \widehat{P}_{-i}(a_{-i,t}|x_t, s_t, z_t, \widehat{\theta}_1))}{\sum_{a_i'} \exp(\sum_{a_{-i}} f_i(a_i', a_{-i,t}, s_{i,t}, x_t; \theta_2) \widehat{P}_{-i}(a_{-i,t}|x_t, s_t, z_t, \widehat{\theta}_1))} \right) \quad (12)$$

The log likelihood (12) takes as given the fitted first-stage estimates of the choice probabilities $\widehat{P}_{-i}(a_{-i,t}|x_t, z_t, \widehat{\theta}_1)$. Given these estimates, we then simply maximize the likelihood of a standard multinomial choice model.

The econometric properties of two-step parametric estimators such as (12) are very well understood. Since $\widehat{P}_{-i}(a_{-i,t}|x_t, z_t, \widehat{\theta}_1)$ will be estimated with error, for any fixed t our estimates will be biased. However, the estimator will be consistent as the number of observations tends to infinity. See Newey and McFadden (1994) section 6 for a succinct discussion of the asymptotic properties of this type of two-step parametric estimator.

22

4.5 Discussion.

The games we study can exhibit multiple equilibria (see McKelvy and Palfrey (1995) and Brock and Durlauf (2001,2003)). An implicit assumption above is that conditional on the observed covariates, x_t, s_t, z_t, there is only a single equilibrium present in the data. The estimation method requires that if we see the same values of x_t, s_t, z_t, the same equilibrium will be selected. This assumption is implicit in most work on estimation of discrete games (see Aguirregabiria and Mira (2003), Pakes, Ovstrovsky, and Berry (2003), Bajari, Benkard and Levin (2003) and Pesendorfer and Schmidt-Dengler (2003)). Recent work has tried to weaken this assumption in similar models (see Sweeting (2004) and Bajari, Hong and Ryan (2004)). However, neither estimator can be applied to the problem that we study.[11]

We acknowledge that this assumption is strong and is unlikely to be perfectly satisfied in any application. However, we believe that it is a useful starting point. Large investment banks have entire departments devoted to research. These departments produce thousands of recommendations each year. The recommendations are publicly observable and actively watched throughout the industry (sometimes recommendations are front page financial news). We argue that the repetition and observability of producing recommendations helps the industry to settle on somewhat standardized norms. Within a single market segment, such as high technology stocks, this makes the single equilibrium assumption more plausible.

At a minimum, our approach contributes to the literature by empirically implementing recent research on the estimation of games. Also, previous approaches to studying recommendations completely abstract away from the simultaneity of analysts' choices. We believe that generalizing the estimation of games to allow for multiplicity will be an active area of future research.

---

[11] The computational burden of Bajari, Hong and Ryan (2004) would be large given the potentially large number of players and strategies in our game. The approach of Sweeting (2004) applied to problems with a dichotomous choice.

Finally, we note that rewriting the utility can give us some further intuition into the identification of the model. In the literature, it is common to consider specifications in which the actions of other players enters into payoffs linearly. Suppose that utility takes the form:

$$u_i(a_i, a_{-i}, x, \varepsilon) = \theta_1 \cdot (x, s_i) + \delta \widehat{P}_{-i}(a_{-i,t}|x_t, s_t, z_t, \widehat{\theta}_1) + \varepsilon_i(a_i) \tag{13}$$

Note that $\widehat{P}_{-i}$ depends on the s_j ($j \neq i$) and z_t. These variables do not directly enter into the rest of the utility function (i.e. (x, s_i)). Our exclusion restrictions therefore guarantee that the first stage estimates are not colinear with the other variables in the utility function. Thus, the model is identified.

5 Empirical Model

In this section, we lay out the model that we will take to the data. An observation is a recommendation submitted for a particular stock during a specific quarter. We will let $q = 1, ..., Q$ denote a quarter, $e = 1, ..., E$ a stock and and $i = 1, ..., I$ an analyst. We will denote a particular recommendation by $r_{i,s,t}$. The recommendation can take on integer values between 1 and 5, where 1 is the highest recommendation and 5 the lowest. Since the dependent variable can be naturally ranked from highest to lowest, we will assume that the utilities come from an ordered probit. While the identification and estimation results from the last section were written for the logit (for expositional clarity), these ideas extend immediately to ordered models.

Let $x_{e,q}$ denote a set of covariates that influence the recommendation of all the analysts for stock e during quarter q. The vector $s_{i,e,q}$ are variables that influence i's recommendation and $z_{e,q}$ are variables that influence the equilibrium selection.

Define the latent variable

$$y_{i,e,q} = \beta' \cdot (x_{e,q}, s_{i,e,q}) + \eta E(r|x_{e,q}, s_{e,q}, z_{e,q}) + \varepsilon_{i,e,q} \tag{14}$$

In equation (14), the term $E(r|x_{e,q}, s_{e,q}, z_{e,q})$ is the expected recommendation for stock e during quarter q and $\varepsilon_{i,s,q}$ is an normal error term. Thus, conforming to the expected actions of peers enters into an individual analyst's utility. In the ordered probit, the probability that a particular recommendation is observed is determined as follows:

$$P(r = 1) = \Phi\left(-\beta' \cdot (x_{e,q}, s_{i,e,q}) - \eta E(r|x_{e,q}, s_{e,q}, z_{e,q})\right) \tag{15}$$

$$P(r = 2) = \Phi(\mu_1 - \beta' \cdot (x_{e,q}, s_{i,e,q}) - \eta E(r|x_{e,q}, s_{e,q}, z_{e,q})) \tag{16}$$

$$-\Phi(-\beta' \cdot (x_{e,q}, s_{i,e,q}) - \eta E(r|x_{e,q}, s_{e,q}, z_{e,q})) \tag{17}$$

$$P(r = 3) = \Phi(\mu_2 - \beta' \cdot (x_{e,q}, s_{i,e,q}) - \eta E(r|x_{e,q}, s_{e,q}, z_{e,q})) \tag{18}$$

$$-\Phi(\mu_1 - \beta' \cdot (x_{e,q}, s_{i,e,q}) - \eta E(r|x_{e,q}, s_{e,q}, z_{e,q})) \tag{19}$$

$$P(r = 4) = \Phi(\mu_3 - \beta' \cdot (x_{e,q}, s_{i,e,q}) - \eta E(r|x_{e,q}, s_{e,q}, z_{e,q})) \tag{20}$$

$$-\Phi(\mu_2 - \beta' \cdot (x_{e,q}, s_{i,e,q}) - \eta E(r|x_{e,q}, s_{e,q}, z_{e,q})) \tag{21}$$

$$P(r = 5) = 1 - \Phi(\mu_3 - \beta' \cdot (x_{e,q}, s_{i,e,q}) - \eta E(r|x_{e,q}, s_{e,q}, z_{e,q})) \tag{22}$$

In equations (14)-(22), the likelihood that determines the probability that the recommendation is r depends on the parameters β and η along with the cut points $\mu_1 - \mu_3$.

5.1 Exclusion Restrictions

The analysis of the previous suggestion suggests that identification depends crucially on having appropriate exclusion restrictions. First, as in A3, we can use covariates that influence the payoffs of one particular agent, but not other agents. In our analysis, the covariates will include broker fixed effects and covariates that reflect broker specific levels of investment banking activity. This assumption would imply, for instance, that the investment banking level of Merrill Lynch should not directly influence the recommendations submitted by Goldman Sachs. We believe that this is a reasonable assumption.

Second, as in A5, identification can be achieved by using covariates $z(e, q)$ that influence the selection of equilibrium, but which do not directly enter into payoffs. The first set of covariates is based on the actions of market regulators. Beginning in June of 2001, the State Attorney General of New York, Elliot Spitzer, began to question business practices in this industry. Spitzer criticized investment banks for issuing a large fraction of strong buy and buy recommendations, but few hold or sell recommendations. The descriptive analysis in section 3 suggests that this intervention by the regulator encouraged the industry to focus on an equilibrium where more "sell" and "hold" recommendations were issued. We divide time into pre and post "Spitzer" eras with the mid point being the quarter starting in June 2001. Based on coverage in the Wall Street Journal, we believe this is an appropriate starting point for the very vocal and public criticism of the industry by a regulator. Therefore, a first instrument is this dummy variable. We believe that the actions of Spitzer helped to shift the equilibrium to a more conservative set of recommendations. It is plausible to argue that SPITDUM can be excluded from payoffs. The analysts could scarcely be accused of conflicts of interest for recommendations issued at this time because there was very little investment banking work being done after June of 2001 due to the downturn high technology markets.

The variable SPITDUM only displays time-series variation. We will interact SPITDUM with SBANK in order to get cross sectional variation in the instrument as well. This would be an appropriate instrument if the criticism of the regulators had more impact for stocks where a larger proportion of the analysts were affiliated with brokerages that had traditionally done investment banking work with high technology firms.

A second set of variables that could enter z include average recommendations for the stock submitted in previous quarters. Table 3.3 demonstrated that in the first quarter of 2000, there were no holds or sells issued by any analyst in our data set. If an individual analyst deviated from this practice, he might miscommunicate to investors his views about the relative desirability of stock e. Therefore, we include behavior in the most recent period since this will communicate information about the norm. Normally, we would be concerned that this instrument is itself arguably endogenous. However, we note that we can control for time and stock effects since we have a panel data set. Also, since we also have access to the instruments based on the actions of the regulator, we can check the robustness of our results both with and without using lagged behavior as an instrument.

6 Results

In this section, we discuss the results from our empirical analysis. We will organize our discussion parallel to the theories discussed in section 3. In section 6.1 we run an ordered probit to study the influence of fundamentals such as earnings forecasts and time and stock fixed effects on recommendations. In section 6.2 we run an ordered probit that includes our measures of conflicts of interest. Finally, in section 6.3, we consider the full-blown model that allows for peer effects in the estimates. The models in 6.1 and 6.2 are single agent ordered probits.

6.1 Fundamentals

The first question that we ask is the extent to which recommendations were determined by publicly observable information about the stocks. In our data, these fundamentals correspond to time fixed effects, stock fixed effects, and the difference between an individual analyst's beliefs about earnings and beliefs in the market as whole. In Table 6.1, we run an ordered probit to explore these questions. The variable %DEV is the percentage deviation of an analyst's recommendation from the average recommendation and ABS. DEV is the algebraic difference. In both cases, a more optimistic earnings forecast has the anticipated sign (-), but is not significant at conventional levels in any of the specifications that we have tried. On the other hand, quarterly and stock fixed effects are almost all statistically significant. If quarter and stock fixed effects proxy for publicly available information about the stock, then this information is considerably more important than measures of an individual analyst's optimism. We note, however, that there could be other interpretations of these variables.

In Table 6.2, we plot the estimated quarterly fixed effects against the NASDAQ Index and NASDAQ 100. The quarterly effects correspond to the model estimated in the last column of Table 6.1. The quarterly effects are labeled qdum2-qdum22 for the 2nd through the 22nd quarter of our data set. Several points emerge. First, the quarterly effects are typically significant. Second, the quarterly effects are highly correlated with the NASDAQ index and with the QQQ. We take these results as evidence that the quarterly effects can reasonably be interpreted as reflecting publicly observed information about firm value as opposed to some other latent effects. The movement in share prices can explain most of the movement in the recommendations in the previous tables. In Table 6.3, we regress the quarterly dummies on these indexes. We find that the indexes are statistically significant at conventional levels and that the measures of

28

goodness of fit are quite high.

6.2 Conflicts of Interest

In table 6.4 we run an ordered probit on recommendations as a function of our conflict of interest measures. The coefficient on RELATION indicates that potential conflicts of interest are statistically significant at conventional levels, except for the third column. The coefficient's sign is also consistent with our a priori beliefs that conflicts of interest could lead to issuing more favorable recommendations. However, these results must be interpreted with some caution. Since brokerage firms are expected to cover companies with whom they have significant investment banking business, the firms have an incentive to select brokerages that already view them favorably. It would be hard to imagine that a rational manager would want to hire an investment banking firm that views her company in an unfavorable manner.

Interestingly, we note (column 4 of table 6.4) that when we include the IBANK variable denoting that the brokerage firm has investment banking relationships with at least one firm in the index, the coefficient is positive and significant. Evidently, analysts at general investment banking firms tend to issue slightly more conservative recommendations. This is consistent with a view of the world that companies tended to select investment banking firms that were more favorably disposed towards them. Alternatively, our results could be interpreted as suffering from a bias due to some omitted variable.

Our results suggest that even though investment banking relationships may generate potential conflicts of interest for equity analysts, the magnitude of the effects recommendations appear to be small. Notice that measures of the goodness of fit are very low when only investment banking is included. Also, the marginal effects on recommendations of engaging in investment banking are small. In our data, the expected recommendation is -0.08 when we allow for quarterly fixed effects, but not stock fixed effects. This finding

29

is not consistent with the prosecutors' belief that "unbiased" research, separate from investment banking, will generate recommendations significantly less tainted by potential conflicts of interest. However, the behavior of analysts after large settlements have been paid and significant damage has been done to their brokerage's reputation, in some cases, may be more conservative.

6.3 Peer Effects

The final question we consider is whether there are peer effects in submitting recommendations. We do this by using the two-stage procedure described in the previous section. First, we flexibly model the expectations of individual analysts about the recommendations that will be submitted by other analysts. We begin by regressing the average recommendation submitted for a particular stock within a given quarter on the following covariates:

1. time dummies
2. stock dummies
3. SBANK- the fraction of analysts that have IBANK equal to one
4. SPITDUM- a dummy for whether the quarter follows the quarter in which Elliot Spitzer began publicly questioning the research quality of Wall Street analysts
5. ESHIFT-the interaction of spitdum and sbank
6. LAG-the lagged value of the average recommendation for the stock

In our analysis, we will assume that variables 4-6 above are valid exclusion restrictions in the sense that they shift the equilibrium that is played in the market, but do not directly enter into an analyst's utility function. The excluded variables are significant in this first-stage regression. We will let IVBELIEF denote the fitted value of the regression and include this variable in the second stage as measure of analysts' beliefs about their peers.

The results from the second-stage are in Table 6.5. In all of the specifications that we examine, peer effects seem to be important. An individual analyst will raise his recommendation proportionally to the

recommendation that he expects from other analysts. This is intuitive. A given recommendation does not make sense in isolation, but only relative to the recommendations of other analysts. For example, if all analysts consider a stock "average" in some sense, then an individual analyst would give the wrong signal by issuing a recommendation of "3" when his peers issue a recommendation of "2".

It is worth noting that our peer effect results are not only statistically significant. Peer effects also explain the results quite well compared to the other covariates. The Pseudo-R^2 using the average recommendation alone is 0.07. Including time and stock effects (our measure of fundamentals) raises it to 0.085. The investment banking relationship is not significant in our final column and does not increase measures of goodness of fit.

7 Conclusion

Two factors seem to be most important for explaining the production of stock recommendations. First, publicly observable information about the stocks under recommendation, as reflected in our time and quarter dummies, plays a large role in explaining the distribution of recommendations. As we saw in Table 4.3, these variables explained a large fraction of the variation in the data and were highly correlated with market indexes such as the NASDAQ or QQQ. Simply put, recommendations improved in 1999-2000 as the stock market rose. The second and most important factor for explaining recommendations is the peer group effect. Individual analysts raise their recommendations proportionally to the recommendations they expect from their peers. Investment banking relationships are shown to be statistically significant in the recommendations regressions, but the economic effect of this variable is estimated to be small. Indeed, this variable is not statistically significant at all in the full-blown model with peer effects.

In the wake of the many conflict of interest allegations that surfaced following the collapse of technol-

31

ogy shares in 2000, policymakers have proposed a wide range of reforms for the industry. These proposals range from mandating greater separation of the investment banking and the brokerage functions, to forcing the complete spin-off of research from brokerage firms. Taken at face value, our findings suggest that none of these proposals would have much effect on reducing conflicts of interest. If more stringent regulation imposed costs on brokerage firms, this could discourage spending on research and limit the flow of information to the markets. Our sample of recommendations on technology firms is taken from a period when the potential for investment banking conflict of interest was relatively great, yet we still find little impact of the investment banking relationship on recommendations.

8 References.

[1] Aguirregabiria, V., and Mira, P., 2002. "Sequential Simulated-Based Estimation of Dynamic Discrete Games." Boston University Working Paper.

[2] Bajari, P., Benkard, C.L., and Levin, J., 2003. "Estimating Dynamic Models of Imperfect Competition." Stanford and Duke University Working Paper.

[3] Bajari, P., Hong, H., and Ryan, S., 2004. "Identification and Estimation of Discrete Games of Complete Information." Duke University Working paper.

[4] Barber, B., Lehavey, R., McNichols, M., and Trueman, B., 2003. "Reassessing the Returns To Analysts' Stock Recommendations." March/April, *Financial Analysts Journal*.

[5] Barber, B., Lehavey, R., McNichols, M., and Trueman, B., 2001. "Can Investors Profit from the Prophets? Security Analyst Recommendations and Stock Returns." *Journal of Finance*. 56, 531-63.

[6] Berry, S., 1992. "Estimation of a Model of Entry in the Airline Industry." *Econometrica*, 60, 889–917.

[7] Bresnahan, T., and Reiss, P., 1990. "Entry in Monopoly Markets." *Review of Economic Studies*, 57, 531–553.

[8] Bresnahan, T., and Reiss, P., 1991. "Empirical Models of Discrete Games." *Journal of Econometrics*, 48, 57-81.

[9] Brock, W. and Dulauf, S. (2001): "Discrete Choice with Social Interactions," *Review of Economic Studies*, 62(2), 235-260.

[10] Brock, W. and Durlauf, S. (2003): "Multinomial Choice with Social Interactions," University of Wisconsin Working Paper.

[11] Chan, L., Karceski, J., and Lakonishok, J., 2003. "Analysts' Conflict of Interest and Biases in Earnings Forecasts." University of Illinois working paper.

[12] Cowles, A., 1933. "Can Stock Market Forecasters Forecast?" *Econometrica*. 1, 309-324.

[13] Dugar, A., and Nathan, S., 1995. "The Effect of Investment Banking Relationships on Financial Analysts' Earnings Forecasts and Investment Recommendations." *Contemporary Accounting Research*. 12, 131-160.

[14] Francis, J. and Philbrick, D., 1993. "Analysts' Decisions as Products of a Multi-task Environment." *Journal of Accounting Research*. 31, 216-230.

[15] Grossman, S., and Stiglitz, J., 1980. "On the Impossibility of Informationally Efficient Markets." *The American Economic Review*. 70, 393-408.

[16] Heckman, J. 1990. "Varieties of Selection Bias." *American Economic Review*. 80(2), 313-318.

[17] Hotz, V., and Miller, R., 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *Review of Economic Studies*. 60,497-531.

[18] Korajczyk, R., Lucas, D., and McDonald, R., 1991. "The Effect of Information Releases on the Pricing and Timing of Equity Issues." *Review of Financial Studies*. 685-708.

[19] Lin, H., and McNichols M., 1998. "Underwriting Relationships, Analysts' Earnings Forecasts, and Investment Recommendations." *Journal of Acounting and Economics*. 25, 1-34.

[20] McKelvy R., and and Palfrey T., 1995. "Quantal Response Equilibria for Normal Form Games." Games and Economic Behavior. 10, 6-38.

[21] Michaely, R., and Womack, K., 1999. "Conflict of Interest and Credibility of Underwriter Analyst Recommendations." *Review of Financial Studies*. 12, 653-686.

[22] Pakes, A., Ovstrovsky, M., and Berry. S., 2003. "Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry / Exit Examples), Harvard University Working Paper.

[23] Pesendorfer, M., and Schmidt-Dengler, P., 2003. "Identification and Estimation of Dynamic Games." London School of Economics Working Paper.

[24] Seim, K., 2001. "Spatial Differentiation and Market Structure: The Video Retail Industry." Yale University Ph.D. Dissertation.

[25] Sweeting, A., 2004. "Coordination Games, Multiple Equilibria and the Timing of Radio Commercials." MIT Working Paper.

[26] Tamer, E., 2002. "Incomplete Bivariate Discrete Response Model with Multiple Equilibria." *Review of Economic Studies*. 70, 147–167.

[27] Womack, K., 1996. "Do Brokerage Analyst Recommendations Have Investment Value?" *Journal of Finance*. 51(1), 137-167.

Table 3.1: Recommendation Variables.

Recommendation	Numerical Value Recorded by I/B/E/S
Strong Buy	1
Buy	2
Hold	3
Underperform	4
Sell	5

Table 3.2: Summary Statistics.

Variable	Mean	Std.	Min.	Max.	Nobs
Recommendation	2.210	0.9168	1	5	12719
Relation	0.0350	0.1839	0	1	12719
Ibank	0.8155	0.3878	0	1	12719
Earnings	0.1111	0.2439	-3.010	1.720	12719

Table 3.3: Tabulation of Recommendations by Quarter.

Variable/Time Period	Q1 1998	Q1 2000	Q2 2003
% Recs. Equal to 1	30.51	46.73	11.65
% Recs. Equal to 2	30.51	41.46	18.12
% Recs. Equal to 3	37.62	11.81	53.07
% Recs. Equal to 4	1.02	0.00	12.62
% Recs. Equal to 5	0.34	0.00	4.53

Table 6.1 Ordered Probit Estimates of the Effect of Fundamentals.

Varible	Coef.	Coef.	Coef.	Coef.
%DEV	-.0539 (-0.276)	-.1030 (-0.519)	-	-
ABS. DEV			-.1030 (-0.519)	-
Log Likelihood	-16171.589	-14861.218	-14861.218	-14861.352
Psueo-R^2	0.0000	0.0810	0.0810	0.0810
Fixed Effects	none	quarterly,stock	quarterly, stock	quarterly, stock

In the ordered probit model, the dependent variable is the analyst's recommendation as coded by IBES. This takes on discrete values from one to five. In the table above, t-statistics are included in parentheses. Most of the quarterly and stock fixed effects are significant in the specifications that we study.

Table 6.2 Quarterly Effects Versus Market Indexes.

Variable	Coefficient	T-Statistic	NASDAQ	QQQ
qdum2	-0.1865	-1.979	1,770	
qdum3	-0.15266	-1.803	1,509	
qdum4	-0.24117	-2.258	1,928	
qdum5	-0.27011	-3.134	2,207	102.25
qdum6	-0.36868	-4.297	2,467	103.87
qdum7	-0.46412	-5.485	2,752	120.12
qdum8	-0.43603	-5.149	3,341	148.63
qdum9	-0.6408	-7.394	4,732	214.5
qdum10	-0.43113	-4.973	3,471	85.19
qdum11	-0.35704	-3.958	4,252	103
qdum12	-0.08381	-1.031	2,664	64.06
qdum13	0.07177	0.955	2,126	46.97
qdum14	0.04712	0.632	2,131	85.19
qdum15	0.063716	0.79	1,802	36.51
qdum16	0.128908	1.671	1,915	39.29
qdum17	0.217669	2.86	1,745	34.15
qdum18	0.349657	4.631	1,613	30
qdum19	0.630952	8.753	1,403	26.05
qdum20	1.009163	13.389	1,887	21.07
qdum21	0.572553	7.596	1,345	24.72
qdum22	0.989535	12.45	1,374	26.06

Table 6.3 Regression of Dummies on Market Indexes.

Variable	Coefficient	Coefficient
Constant	.8208896 (3.965)	0.6270 (4.3)
Nasdaq Index	-.0003467 (-4.960)	-
QQQ Price	-	-0.007 (-4.7)
Nobs	21	18
R^2	0.48	0.6830

Table 6.4 Ordered Probit Estimates of the Effect of Conflicts of Interest.

Varible	Coef.	Coef.	Coef.	Coef.
RELATION	-.3231 (-6.19)	-.1108 (-2.06)	.05397 (0.94)	.03932 (0.68)
IBANK				.1080 (4.18)
Log Likelihood	-16152.389	-15314.605	-14860.888	-14855.94
Psueo-R^2	0.0012	0.0530	0.0811	0.0814
Fixed Effects	none	quarterly	quarterly, stock	quarterly, stock

In the ordered probit model, the dependent variable is the analyst's recommendation as coded by IBES. This takes on discrete values from one to five. In the table above, t-statistics are included in parentheses. We do not report ancillary parameters, such as the cut values and values of the fixed effects.

Table 6.5 Ordered Probit Estimates including Peer Effects.

Varible	Coef.	Coef.	Coef.	Coef.
IVBELIEF	1.289 (46.16)	1.2911 (22.174)	.7603 (7.684)	.7816 (7.47)
RELATION				.0537 (0.93)
%DEV				-.1140 (-0.57)
Log likelihood	-14720.112	-14719.611	-14475.708	-14476.514
Psuedo-R^2	0.0695	0.0695	0.0850	0.0849
Fixed Effects	none	stock	quarterly, stock	quarterly, stock

In the ordered probit model, the dependent variable is the analyst's recommendation as coded by IBES. This takes on discrete values from one to five. In the table above, t-statistics are included in parentheses. Most of the quarterly and stock fixed effects are significant in the specifications that we study.