

## On the Structure of Analyst Research Portfolios and Forecast Accuracy

Omesh Kini  
Robinson College of Business  
Georgia State University  
e-mail: okini@gsu.edu

Shehzad Mian  
Goizueta Business School  
Emory University  
e-mail: shehzad\_mian@bus.emory.edu

Michael Rebello  
A.B. Freeman School of Business  
Tulane University  
e-mail: mrebello@tulane.edu

Anand Venkateswaran  
College of Business  
Northeastern University  
e-mail: anand@neu.edu

Latest Draft: November 2006

We would like to thank Sudipta Basu, George Benston, Lawrence D. Brown, Diane Denis, Karl Lins, Mike Lemmon, Daphne Lui, Stan Markov, Grace Pownall, Raghu Rau, Jeremy Stein, and Gregory Waymire. Discussions with Ashish Bhutani (Lazard), Mathew Chapin (UBS Warburg), Devashish Chopra (formerly with Citi-SSB Asset Management), and Dan Meade (formerly with Credit Suisse First Boston) have given us valuable insights from a practitioner's perspective. We thank *IBES* for providing us permission to use their database and Ajay Negi of *IBES* for patiently answering all our questions regarding *IBES International*. We acknowledge valuable research support from Terry Bennett and Ron Harris. The usual disclaimer applies.

## On the Structure of Analyst Research Portfolios and Forecast Accuracy

### ABSTRACT

This paper provides insights into the determinants of the structure of analyst research portfolios, and then investigates the impact of portfolio structure on forecast accuracy. Specifically, we study the forces and constraints that shape analyst research coverage along country and sector dimensions. We find that analyst specialization is sensitive to the extent to which firms within a country or sector are exposed to common economic forces, the potential for revenue generation, and broker culture. Further, in contrast to the existing research on the relationship between analyst portfolio structure and forecast accuracy, we find that complex research portfolios promote forecast accuracy once we control for the endogeneity in portfolio choice. Overall, our results are consistent with brokerages organizing analyst research coverage in a pragmatic manner. While maintaining a focus on the economic viability of the research production process, the organization of brokerage research is consistent with producing high quality research.

*JEL Classification:* G24; G15; L23

*Keywords:* International Analyst Coverage, Brokerage Organization, Forecast Accuracy, Capital Markets

## **On the Structure of Analyst Research Portfolios and Forecast Accuracy**

Recent research has documented that U.S. analysts frequently cover stocks from multiple sectors, and international analysts often follow stocks belonging to both different countries and sectors (see, e.g., Clement and Tse (2005), Boni and Womack (2005), and Sonney (2005)). Recognizing that this diversity in research portfolios can affect analysts' research output, researchers have attempted to assess its effect on analyst performance as captured by forecast accuracy (see, e.g., Clement (1999), and Clement, Rees, and Swanson (2003)). Existing studies, however, have treated the structure of analysts' research portfolios as an exogenous variable which may cause their results to be subject to a selection-bias given that analyst research portfolios are the result of choices made by analysts and their employing brokerages (see, e.g., Clement (1999), and Gilson, Healy, Noe, and Palepu (2001)).

In this paper, we examine the impact of portfolio structure on analyst forecast accuracy while controlling for selection biases resulting from portfolio choices made by analysts/brokerages. To accomplish this, we first examine the determinants of one aspect of analyst organization—the choice between either single- versus multi-country coverage or single- versus multi-sector coverage. We then use this analysis as a means of controlling for self-selection in our examination of the relation between forecast accuracy and portfolio diversity in the country and sector dimensions.

We employ a comprehensive sample of international (non-U.S.) analysts from 53 countries over the period 1996 through 2002. We focus on international analysts for the following reasons. First, there is limited evidence on both the determinants of research portfolio structure and forecast accuracy for international analysts. This omission in the literature becomes all the more glaring once we consider the fact that international analysts account for

approximately 60% of all analysts in the *IBES* database. Second, by using international analysts we add a dimension that would be missing in an analysis of only U.S. analysts since analysts in our sample can diversify by following firms either across countries or sectors.

To examine the choice of single- versus multi-country (single- versus multi-sector) coverage, we estimate binomial logistic regressions to explain specialization of analyst portfolios by country (sector). We find that increased commonalities among stocks in a country, as proxied by common variation in returns, increase the likelihood that analysts will specialize by country. Similarly, increased commonalities among stocks in a sector promote sector specialization. We also find that higher country equity market capitalization increases the likelihood of country specialization, while a larger total equity market capitalization of a sector in a particular country (country-sector market capitalization) increases the likelihood that analysts will cover firms in a single sector. Finally, we find that international analysts employed by a U.S. brokerage are more likely to specialize by sector and less likely to specialize by country.

We then consider the impact of portfolio choice on forecast accuracy. To establish a benchmark, we first follow the extant literature and treat portfolio structure as an exogenous variable in our tests. We find that neither country diversification nor sector diversification has an effect on forecast accuracy. These results are similar to those documented by Clement, Rees, and Swanson (2003) in their investigation of Canadian, German, Japanese, and U.K. analysts. Next, we investigate the relation between forecast accuracy and portfolio organization by employing a two-stage Heckman methodology that controls for research portfolio choice.<sup>1</sup> The results from this analysis support our central thesis that self-selection can have a significant

---

<sup>1</sup> See, for example, Heckman (1978; 1979), Heckman and Robb (1986), and Greene (2000) for a general approach to estimate a causal parameter in the face of selection bias. Also see, Li and Prabhala (2005) for a survey on the use of self-selection models in finance research.

impact on analyst forecast accuracy. Further, once we employ the Heckman procedure to control for endogeneity in analyst portfolio choices, we find that diversified portfolios boost analysts' forecast accuracy. This is the case whether we consider portfolio diversification across either countries or sectors. These findings stand in contrast to the results in earlier studies which document that, when analyst organization is treated as an exogenous variable, portfolio diversification either has no impact or detracts from forecast accuracy (see, e.g., such as those of Clement (1999), Clement, Rees, and Swanson (2003), and Clement and Tse (2005)).

Our explanation for these results is as follows: A focused portfolio should enable an analyst to enjoy economies of scale in information acquisition and production. The time and effort savings resulting from these scale economies could be transferred to gaining a more thorough understanding of the companies she follows and, thus, enable the analyst to construct more accurate forecasts. By switching to a diversified portfolio, an analyst will lose some scale economies but gain exposure to complementary information that is needed to research companies in other countries or sectors. By accessing and analyzing this complementary information, the analyst may obtain a more complete picture of the economic forces at play, and thereby generate more accurate forecasts. The benefit from accessing complementary information is likely to dominate the resulting loss in scale economies when the possibility of harnessing scale economies by maintaining a focused portfolio is relatively low. Thus, analysts will tend to follow diversified portfolios when the cost of diversification, i.e., the loss of scale economies, is relatively low. This is precisely what our results seem to suggest as (i) the diversified portfolios in our sample boost forecast accuracy and (ii) analysts tend to follow diversified portfolios precisely when the common variation in returns, our proxy for the presence of scale economies, is relatively low for stocks belonging to the same country (sector).

Our results provide several new insights into the organization structure of brokerages and the influence of this structure on analyst performance. First, our results support the notion that brokerages organize their research production in a pragmatic fashion. Consistent with revenue generation being essential for the existence of sell-side research, we find that the decision to specialize by country or sector is determined by the potential for investment banking business and/or brokerage commissions.<sup>2</sup> However, we also find that brokerages structure their research production processes in a manner that is conducive to producing quality research by allowing analysts to follow diversified portfolios and consequently sacrifice the economies of scale in information acquisition and production resulting from this diversification precisely when the cost of the lost scale economies is relatively low. This ensures that, on average, diversified research portfolios outperform focused research portfolios.

Our paper relates to and has implications for several streams of research. It is closely related to the literature that examines how the organizational structures of financial intermediaries evolve in response to their information processing environment. For instance, Berger et al. (2005) find evidence that the nature and availability of information regarding firms leads large banks to focus on lending to large firms, while smaller banks target smaller firms. They attribute this difference in focus across banks of differing sizes to the cross-sectional relationship between the mix of hard and soft information and firm size. In a similar vein, we document that specialization of sell-side analyst research portfolios is influenced by information availability on the firms that analysts research. When efficiency in gathering and processing country-specific (sector-specific) information is increasingly vital, analysts tend to focus their coverage on an individual country (sector).

---

<sup>2</sup> Chen (2003), for example, argues that the separation of investment banking from sell-side research to solve the conflict of interest problem may make sell-side research infeasible; thereby diminishing information production in the economy and reducing social welfare.

Recent studies have shown that analysts' forecast accuracy is systematically related to the geographical proximity of analysts to the companies they research. For example, Bacmann and Bolliger (2001) provide evidence indicating that foreign analysts provide more accurate forecasts than local analysts. In contrast, Bae, Stulz, and Tan (2005), Bolliger (2001), Malloy (2005), and Chang (2003) document that locally based analysts make more accurate forecasts than their counterparts who are based abroad.<sup>3</sup> Our analysis complements this literature as it focuses on the "proximity", as measured by the extent of exposure to common economic factors, of the companies covered by the analysts to each other rather than the geographical proximity of the analysts to the companies they cover. This difference in perspective provides us with a rationale that can potentially reconcile the conflicting results in the above cited studies. For example, if a good understanding of sector factors is crucial to forecasting accurately, a foreign analyst who focuses on a sector may have an advantage over a local analyst who focuses on local firms across multiple sectors. The reverse would be true if an understanding of country (local) factors is the key determinant of forecast accuracy. This suggests that comparisons of the performance of local and foreign analysts should control for differences in the industrial structure among countries.

Finally, and most importantly, we directly contribute to the literature examining the role of analyst portfolio complexity on forecast accuracy by examining the forces and constraints that drive analysts to select complex portfolios (see, for example, Clement (1999) and Clement, Rees, and Swanson (2003)). In doing so, we provide insight into the cross-sectional variation in analysts' costs and benefits from diversifying across sectors and countries. Further, our analysis

---

<sup>3</sup> There exists a parallel literature on geography and investment performance. Kang and Stulz (1997), Choe, Kho, and Stulz (2000), and Coval and Moskowitz (2001) find that proximity improves performance, while Grinblatt and Keloharju (2000) and Seasholes (2004) document the opposite result.

highlights the need to control for self-selection in portfolio choice when studying the relation between forecast accuracy and analyst portfolio structure.

The remainder of the paper is organized as follows: Section I provides the underlying hypotheses that are tested in this paper. Section II contains a description of our data sources and sample selection screens. Section III contains our analysis of the determinants of analysts' portfolio choice along both country sector dimensions. In Section IV, we examine the impact of portfolio complexity on forecast accuracy. Section V presents results of tests designed to provide further insight into our central results and to assess their robustness. The paper concludes with some final observations in Section VI.

### **I. Development of Testable Hypotheses**

In this section, we first develop hypotheses related to the determinants of analyst organization. The predicted relations with analyst organization are summarized in Table I. Subsequently, we develop hypotheses relating analyst organization to performance as measured by the accuracy of analysts' current fiscal year-end forecasts.

#### *I.A. Determinants of Analysts Research Portfolio Structure*

*Information Efficiency Hypothesis.* Sets of firms operate in common markets and, thus, common forces influence their costs and revenues. These commonalities across firms allow an analyst who has studied a firm to expend less time and effort to determine the value of other firms exposed to the same economic forces. Thus, by focusing her attention on a set of firms that are strongly influenced by a common set of forces, an analyst is able to harness economies of scale in the acquisition and production of information. By taking advantage of these scale economies, an analyst may be able to study firms in greater depth and thus produce more accurate earnings forecasts.



Less focused portfolios may also promote forecast accuracy because a less focused portfolio, while resulting in the loss of scale economies in information acquisition and production, may expose the analyst to alternate sources and types of (complementary) information regarding the firms she covers. For example, an analyst covering utility companies may benefit from the information and knowledge gained from the coverage of the oil and natural gas industry. Similarly, an analyst covering the ship manufacturing industry may be better positioned to assess the size of the potential order books of ship manufacturers if she also covers firms in the shipping industry.

The extent of focus in analyst portfolios is likely to be influenced by this tradeoff between the relative information-based benefits of focus and diversity. At the margin, however, an increase in commonalities will strengthen the incentives for an analyst to specialize. Thus, we hypothesize that an analyst is more likely to confine her research portfolio to a set of firms as the strength of commonalities across the set of firms increases. We focus on two sources of commonality—shared country of domicile and shared industrial sector membership—and hypothesize that an analyst will tend to pick a country-focused portfolio as country-based commonalities increase. Similarly, as sector-based commonalities increase, the likelihood of a sector-focused portfolio improves. In our analysis, we decompose equity market returns in a manner similar to that in Heston and Rouwenhorst (1994) and Griffin and Karolyi (1998) to estimate the extent of commonalities between firms belonging to the same country (*VARCTYTOT* or *VARCTYSYS*) or sector (*VARSECTOT* or *VARSECSYS*). A description of the construction of these variables is presented in the next section.

*Revenue Generation Hypothesis.* Market size may also influence the structure of an analyst's research portfolio. Analysts have to ensure that the information they produce generates sufficient revenues for their employer. These revenues take the form of brokering commissions

and investment banking fees, both of which are a function of the size of an analyst's portfolio. Thus, a smaller market is likely to increase the cost of focusing coverage on the market and, thus, discourage an analyst from restricting her entire coverage to the market. Thus, we hypothesize that specialization is more likely when market size is larger.<sup>4</sup>

In our analysis, the natural logarithm of equity market capitalization (*CTYMV*) serves as a proxy for market size. The likelihood of country specialization, therefore, will increase with the country's equity market capitalization. Global sector market capitalizations are large and, thus, are unlikely to be binding constraints on portfolio choice. Sector capitalizations within a country are, however, relatively small, and are likely to act as constraints on sector specialization for analysts who choose to restrict their coverage of firms to a single country. Thus, the likelihood of sector specialization will increase with the natural logarithm of the market capitalization of a sector within a country, *CTYSECMV*.

*Broker Culture Hypothesis.* It is argued that there are two well-established and distinct research models—the European and U.S. models—that have evolved in the context of very different domestic markets (see, e.g., Rubino (2003)). U.S. broking firms are reputed to have attempted to export their sector-focused organization structure as they have expanded internationally, encouraging their analysts to focus on a single sector. Conversely, European firms are reputed to lean toward a country-focused organization, biasing their analysts towards focusing their research coverage on a single country (See, for example, *Reuters Institutional Investor Research Group* (2002)). Analyst research portfolios are likely to reflect the organizational culture and resources of their employing brokers. Therefore, we hypothesize that an analyst's research focus is influenced by her employer's traditional research organizational structure. Analysts

---

<sup>4</sup> This reasoning echoes Stigler (1951) who identifies the extent of the market as a limiting factor on specialization.

employed by European brokerages will display a greater tendency to specialize by country, while analysts employed by U.S. brokerages are more likely to specialize by sector. In our analysis, we use a broker's domicile as a proxy for its research culture and traditional organization structure. More specifically, we employ a dummy variable, *BRORIGIN* that takes the value of one if the broker is a U.S broker, and zero otherwise.

Brokers can also vary in the opportunities that they afford an analyst to specialize. We hypothesize that the opportunities for specialization are positively related to broker size.<sup>5</sup> Larger brokers cover a larger number of stocks and are, therefore, less likely to constrain an analyst portfolio away from specialization along any given dimension. We use the number of analysts employed by the brokerage firm, *BRSIZE*, as a proxy for broker size.

*Analyst Experience Hypothesis.* Just as a pecking order theory of capital structure argues that a firm's capital structure is the culmination of its history of security issuance choices, analyst portfolios may be the result of a history of portfolio decisions. Experienced analysts are more likely to have changed employers and their past employers may have had different research cultures. As a result, they are more likely to maintain research portfolios that contain diverse stocks. Further, more experienced analysts by definition have been in the industry for a longer period of time, and, as such, are likely to have demonstrated superior ability in the past to be still counted amongst the survivors in the industry.<sup>6</sup> Since following stocks across sectors and/or countries may require the ability to conduct complex analyses, brokerages will assign diverse research portfolios to these analysts in keeping with their higher perceived ability.

---

<sup>5</sup> Larger professional firms, such as auditors and law firms, are often associated with more specialization in their practice (see, for example, DeAngelo (1981), Garicano and Hubbard (2005), and Zhou and Elder (2003))

<sup>6</sup> Consistent with this line of argument, Hong and Kubik (2003) and Clement and Tse (2005) find that U.S. analysts who make more accurate earnings forecasts face better employment outcomes. Kini et al. (2006) document similar results for international analysts.

Further, if more experienced analysts indeed have greater ability then we should also observe a positive relation between forecast accuracy and analyst experience. Our proxy for analyst experience, *ANALEXP*, is the number of days between when an analyst first appears in *IBES* until her last forecast for a given fiscal year.

### *I.B. Portfolio Structure and Forecast Accuracy*

As we have argued earlier, the structure of an analyst's portfolio is likely to impact her performance as measured by the accuracy of her current fiscal year-end earnings forecasts. A focused portfolio should enable an analyst to enjoy economies of scale in information acquisition and production, which in turn, can enable the analyst to make more accurate forecasts. Alternatively, by diversifying her portfolio, an analyst gains exposure to complementary information. By incorporating this complementary information in earnings forecasts, the analyst may be able to improve her accuracy. Since the economies of scale and informational complementarities arguments affect forecast accuracy in opposite directions, the predicted relation between forecast accuracy and analyst specialization is ambiguous.

Existing research examines the relation between forecast accuracy and analyst portfolio complexity by estimating regressions that treat the structure of the analysts' portfolios as being exogenously determined. However, the structure of an analyst's portfolio may be the result of choices made by the analyst's employing broker and/or the analyst herself. Further, there may be latent factors that affect both the structure of analyst portfolios and their forecasting ability. For example, it is possible that brokerages prefer to assign diversified portfolios to analysts who are also responsible for supervising other analysts, placing these supervising analysts in a better position to evaluate and monitor the activities of analysts covering different types of firms. These administrative duties, however, may adversely affect the forecasting abilities of the

supervisors. Thus, to the extent that the structure of an analyst's portfolio is determined by choices made by the analyst, tests of the relationship between the structure of an analyst's portfolio and her forecast accuracy should account for the endogeneity of the portfolio structure. In our analysis, we do so by using the Heckman methodology that allows us to account for the self-selection bias.

## II. Description of Data Sources, Analyst Portfolios, and Factor Extraction

### A. Data Sources for Analyst Portfolios and Firm Information

The primary sources of data for this study are *IBES International*, *Datastream International*, and *Nelson's Directory of Investment Research (Nelson's)*. From the *Details* file of *IBES International*, we obtain the list of all analysts and stocks that they cover over our sample period (January 1, 1996 through December 31, 2002).<sup>7</sup> We divide our sample period into one-year intervals for computation of analyst portfolios and other variables. Next, we assign each stock to a unique country and sector pair for each year in our sample in accordance with the *IBES Sector Industry Group* file.<sup>8</sup> We supplement *IBES* with *Nelson's* to identify the employing broker for every

---

<sup>7</sup> Analysts are excluded from our sample if either their name is missing, or *IBES* did not assign them an analyst code, or *IBES* assigned analyst names that are names for industry groups like Health Care or Pharmaceuticals. Our conversations with Mr. Ajay Negi at *IBES* suggest that when broker firms are small or analyst turnover is high, they sometimes report their earnings forecasts using sector names rather than individual analyst names. Another possible explanation is that these types of names are used for a sector or country team.

<sup>8</sup> We deleted firms that belong to the "Miscellaneous/Undesignated" sector in *IBES* because, despite research into these firms, we could not easily classify them into a specific *IBES* sector. In some instances, firms are assigned by *IBES* either to multiple sectors, multiple countries, or both. For firms assigned by *IBES* to multiple countries, we checked other sources to identify the location of the company's headquarters and accordingly assigned each firm to a particular country. We classify firms that *IBES* assigned to multiple sectors to the sector to which *IBES* assigned them for the majority of any given year in our sample period.

analyst.<sup>9</sup> We then construct a research portfolio for every analyst that includes all stocks for which they issued any forecast (quarterly, semi-annual, current-year annual, following-year annual, long-term earnings growth, etc.) during the year in consideration. Our measure of performance is, however, based solely on the accuracy in annual earnings forecasts for the current fiscal year. Analyst forecasts included in our sample correspond to the last forecast issued by an analyst at least one month prior to the fiscal year end, and earnings realizations for each fiscal year for stocks in our sample are obtained from the *IBES* database. Furthermore, we trim the sample at the 1 percentile and 99 percentile levels of raw forecast accuracy. We obtain stock prices, number of shares outstanding, stock returns, and equity market capitalization data from *Datastream International*.<sup>10</sup> The resulting sample consists of 19,379 analysts following 11,292 stocks over the seven-year sample period across 53 countries and 11 sectors, resulting in 47,049 analyst-year observations.

#### *B. Description of Analyst Portfolios*

Table II provides summary statistics for our analyst sample. The mean (median) number of countries followed by an analyst is 1.46 (1.00) indicating that the analysts' coverage appears to be predominantly focused on a single country. Analysts however tend to cover multiple

---

<sup>9</sup> In instances in which *IBES* assigns multiple broker codes to the same broker, we consolidate all broker codes for a given broker using a unique broker code. These unique broker codes are then employed as the basis for all broker-analyst linkages examined in our paper. Even after the multiplicity of broker codes are dealt with, some analysts were assigned to between 2 and 6 brokers by *IBES*. This can either be due to changes in employment or because *IBES* continued linking the analyst with previous employers. To resolve multiple broker associations, we assign analysts to the broker who employed the analyst for the maximum period during a two-year window around the period within which the analyst's affiliation is an issue. Thus, an analyst who appears in the sample for the year 1998 is assigned to the broker with whom he was employed for the longest time between January 1, 1998 and December 31, 1999.

<sup>10</sup> We delete firms from the sample if their average market capitalization was less than zero, if there was no price, market value, or returns data for them during the sample period, or if they were classified in the "Miscellaneous/Undesignated" sector in *IBES*. We also delete analyst-stock-year observations if they had missing forecast accuracy information.

sectors and the mean (median) number of sectors being followed is 2.36 (2.00). The mean (median) number of firms followed by analysts is 8.41 (6.00), and the mean (median) number of days of analyst experience, as measured by the number of days from when the analyst first appears on *IBES* to the time period in which the analyst makes a forecast for a firm in our sample during a given fiscal year, is 841 (606) days.

To gain an understanding of how analysts organize their research portfolios, we develop two classifications for the research coverage of each analyst. First, each analyst is classified either as a single-country analyst or multi-country analyst. Second, each analyst is classified as a single sector or multi-sector analyst. Panels B and C of Table II describe analyst portfolios conditional on the two classifications described above. The vast majority (78.00%) of our observations are for single-country analysts. In contrast, only about 40% of the analysts in our sample cover firms in a single sector. Single-country analysts tend to follow fewer firms (mean value of 8.27 versus 8.91) than multi-country analysts. Analyst experience is approximately four months lower for single country analysts. A similar pattern is apparent when examining analyst portfolios based on the split between single sector and multi-sector coverage—analysts who cover multiple sectors tend to follow more firms and have greater experience.

### *C. Description of Variables Hypothesized to Impact Portfolio Choice*

Note that, while each analyst-stock-year observation in our sample is unique, each analyst enters our sample multiple times based on the number of stocks she follows. This allows for the possibility that regression estimates employing analyst-stock-year observations may be influenced disproportionately by analysts who cover larger portfolios. Consequently, we employ analyst-year observations in the subsequent analysis. To arrive at the sample that is employed to test our hypotheses, for each time period, every variable is averaged across stocks covered by the analyst. For example, if  $x_{ijt}$  is a variable of interest for analyst  $i$  and stock  $j$  in time

period  $t$ , in our analysis we employ the average value,  $x_{it}$ , of this variable across stocks covered by the analyst for time period  $t$ .

These averages of variables across stocks covered by an analyst for a period are taken in two ways. First, we average each variable across all stocks covered by the analyst in a given period. The second method for averaging takes into consideration the fact that, while an analyst may cover a diverse set of stocks, her primary responsibility may be producing research on a set of homogeneous “core” stocks that belong to a single sector within a given country. Thus, we also average across stocks belonging to the “core” of an analyst’s portfolio during that period, where core stocks are stocks that belong to the unique country-sector combination that accounts for the largest portion of the analyst’s research portfolio in that year. All other stocks are considered non-core stocks. For all the results reported in the paper, the core of an analyst’s portfolio is the country-sector that accounts for the largest number of stocks in the analyst’s portfolio (*CORE*). The sectors defined above are based on IBES International’s 11 sectors.<sup>11</sup>

In Table III, we present summary statistics for all the variables employed in modeling analyst organization. The first two rows present statistics for our measures of exposure to common economic forces for firms belonging to the same country, while the next two rows present statistics for our measures of exposure to common economic forces for firms that belong to the same sector. These proxies for the exposure to common economic forces are based on the common variations in stock returns along country (sector) dimensions. Following Heston and Rouwenhorst (1994) and Griffin and Karolyi (1998) we decompose the variance of the weekly returns for each firm  $i$  in country  $j$  and sector  $k$ ,  $\sigma_i^2$ , as follows:

---

<sup>11</sup> In results not reported in the paper, we replicate our entire analysis by defining the core of an analyst’s portfolio as the country-sector that accounts for the largest fraction of the market capitalization of the analyst’s portfolio. Our results are not sensitive to these alternative definitions of an analyst’s core.



$$\sigma_i^2 = \sigma_G^2 + \sigma_{\beta j}^2 + \sigma_{\gamma k}^2 + \sigma_{\epsilon i}^2,$$

where  $\sigma_G^2$  represents the variation in the global factor,  $\sigma_{\beta j}^2$  ( $\sigma_{\gamma k}^2$ ) represents the variation in the country (sector) factor for the country  $j$  (sector  $k$ ) to which firm  $i$  belongs, and  $\sigma_{\epsilon i}^2$  represents the variance of the firm-specific component of returns. The details of the decomposition methodology are presented in Appendix A.

We develop two proxies each for the influence of country and sector factors on firm returns. The proxies for the influence of the country factor on a firm's returns are normalized values of  $\sigma_{\beta j}^2$  for each firm. *VARCTYTOT* is obtained by normalizing the weekly return variation on the country factor by the total weekly firm return variability ( $\sigma_{\beta j}^2 / \sigma_i^2$ ). *VARCTYSYS* is obtained by normalizing the weekly return variation on the country factor by the sum of the variances of the weekly returns on the global, country, and sector factors ( $\sigma_{\beta j}^2 / (\sigma_G^2 + \sigma_{\beta j}^2 + \sigma_{\gamma k}^2)$ ), i.e., the proportion of total "systematic risk" attributable to the country factor. The first row in Table III indicates that, on average, the country factor accounts for 19.9% of total variation in the returns on *core* stocks in an analyst's portfolio. On the other hand, when the average value of *VARCTYTOT* is computed using all stocks in the analyst's portfolio, the country factor accounts for approximately 20.1% of return variation. Similarly, the mean value of *VARCTYSYS* is 55.3% when only stocks belonging to the core of an analyst's portfolio are considered and 55.7% when all stocks in an analyst's portfolio are considered.

The variable *VARSECTOT* (*VARSECSYS*) is the sector counterpart of *VARCTYTOT* (*VARCTYSYS*) in that it measures the proportion of total risk (systematic risk) of firm  $i$  attributable to the sector factor. From Table III, it is clear that the sector factor accounts for a smaller fraction of variation in stock returns than does the country factor. When core stocks in

an analyst's portfolio are considered, the variable *VARSECTOT* has a mean (median) value of 6.6% (4.3%), while *VARSECSYS* has a mean (median) value of 32.0% (19.5%). Both variables have similar mean values when all stocks belonging to the analyst's portfolio are considered.

Every stock in an analyst's portfolio belongs to a particular country. The variable *CTYMV* represents the mean value of the country market capitalization for stocks in the analyst's portfolio. Similarly, *CTYSECMV* represents the mean value of the country-sector market capitalization for stocks in the analyst's portfolio. The figures reported in Table III suggest that market capitalizations are larger for core countries and core country-sectors. The mean value for *BRORIGIN* indicates that 24.3% of analysts in our sample are employed by brokerages of U.S. origin. *BRSIZE* is a count of the number of analysts employed by the broker that employs a given analyst. The mean value of *BRSIZE* is approximately 156 and its median value is 107, suggesting that a significant fraction of the analysts in our sample are employed by relatively large brokers. Our measure of experience, *GENEXP* represents the number of days from when the analyst first appears on *IBES* to the time period in which the analyst makes a forecast for a firm in our sample during a given fiscal year. The mean and median values of this variable suggest that, on average, analysts in our sample have been employed in the profession for approximately two years. Note that the mean (median) values for *BRSIZE*, *GENEXP*, and *BRORIGIN* are the same for all stocks and core stocks in analysts' portfolios since they are measured at the analyst level.

The remaining variables in the table are employed as controls in our examination of analyst portfolios. We now turn to variables employed as controls for the information environment in which analysts operate. *VARRESTOT*, ( $\sigma_a^2 / \sigma_i^2$ ) the proportion of idiosyncratic variation in stock returns, allows us to control for the relative strength of factors other than country and sector factor that might influence the complexity of an analyst's task. *FIRMMV*, the

average market capitalization of an individual stock covered by an analyst in a given year, is employed as a proxy for the disclosure environment. From Table III, it appears that the market capitalizations of firms in the core of an analyst portfolio are similar in size to non-core firms suggesting that the disclosure environment may not vary dramatically between the core and non-core portion of an analyst's portfolio.

*ACCTG INDX*, is an index of country-level disclosure regulation. A higher value for this index indicates greater information disclosure in that country. It is a count of how many of 90 standard accounting items are included in the annual reports of firms in a country. We obtained this information from *International Accounting and Auditing Trends* published by the *Center for International Financial Analysis and Research*. Table III suggests that firms in an analyst's portfolio belong to countries that, on average, require approximately 69 out of the 90 possible accounting items in annual reports and that these disclosure requirements are similar for core and non-core firms.

Improved safeguards for investors should attract investors who initiate trades based on fundamentals (Morck, Yeung, and Yu (2000)), and thereby enhance the value of analyst research.<sup>12</sup> To control for the level of minority shareholder protection in a country, we include a variable *ANTIDIR RIGHTS* in our regressions. The source for this data is La Porta et al. (1997) and the variable represents the average value across an analyst's portfolio of the La Porta et al. variable.<sup>13</sup> The mean (median) value for this variable is approximately 3.4 (4.0) out of a

---

<sup>12</sup> Minority protection will tend to encourage greater ownership dispersion and, therefore, greater capital market liquidity. Enhanced liquidity is likely to generate greater trading volume and higher commission income. This effect will render more focused portfolios economically viable.

<sup>13</sup> *ANTIDIR RIGHTS* indicates how many of the following mechanisms prevail in a given country: (1) shareholders are allowed to mail their proxy vote (2) shareholders are not stipulated to submit their shares before a general shareholders' meeting (3) cumulative voting is allowed (4) there is a system in place to address issues pertaining to oppressed minority shareholders (5) the minimum ownership

maximum of five mechanisms to protect minority shareholder rights for core stocks and all stocks in an analyst's portfolio. The variable, *CTYMVGDP* is a proxy for the degree of capital market development of the country in which the firm is domiciled, and is our final control variable for the level of disclosure. It is the ratio of stock market capitalization to GDP, and has been employed to capture informational availability by other researchers (See, Rajan and Zingales (1998)). The data to compute this variable for each year is obtained from various issues of the *World Stock Exchange Fact Book*. The mean (median) values for *CTYMVGDP* appear to be similar for core and non-core stocks in an analyst's portfolio.

### III. Analyst Organization

In this section, we present results from our logistic regression analysis modeling analyst portfolio choices. First, we examine the determinants of single country versus multiple country research portfolios. We then present results from our logistic regression analysis modeling the determinants of single sector versus multiple sector research portfolios. These tests employ analyst-year level observations. As a consequence, while an analyst appears only once for each time period, an analyst may enter the sample several times over our sample period. Since analysts may appear in the sample multiple times, it opens up the possibility that forecast errors may be correlated and t-statistics may be overstated because of the "cluster sample" problem (see, e.g., Wooldridge (2002)). To control for this issue, we employ adjusted standard errors that account for the possible correlations between forecast errors for an analyst. Our adjustment, however, is made under the assumption that forecast errors across analysts are independent

---

threshold for a shareholder to call an extra-ordinary shareholders' meeting is no greater than ten percent. Thus, this index will lie between zero and five.

(see, e.g., Huber (1967), Rogers (1993), White (1980), and Wooldridge (2002)).<sup>14</sup> In addition, in our analysis of this panel data, we control for year-fixed effects in the estimated regressions.<sup>15</sup>

One way of looking at the choice of an analyst to diversify across countries and/or sectors is to ask the question what factors cause her to follow stocks outside of her core country-sector, i.e., stocks in countries outside her core country or stocks in sectors outside her core sector. We attempt to answer this question by both studying the impact on portfolio choice of (i) the characteristics of just the stocks in her core country-sector and (ii) the characteristics of all the stocks in her research portfolio. Thus, each model of analyst portfolio choice is estimated two ways. One estimate employs independent variables values that are averaged across all stocks in an analyst's core while the second estimate employs independent variable values that are averages across an analyst's entire portfolio.

### *III.A. Analysis of Single versus Multiple Country Coverage*

Table IV presents tests of our hypotheses regarding analyst specialization by country. We employ a binomial logistic regression approach to model the probability of being organized as a multiple-country versus single-country analyst.<sup>16</sup> The first two models reported in the table are identical with the exception of the variable that represents the country factor--*VARCTYTOT* is employed in the first model and *VARCTYSYS* in the second model. All independent variables employed in these two regressions are average values across all stocks in the analyst's core country-sector for a given period. The third and fourth models reported in

---

<sup>14</sup> See Gleason and Lee (2003) for a similar approach to dealing with cluster sample problems. Further, we also employ adjusted standard errors that account for the possible correlations between forecast errors for analysts that work for the same brokerage. Our results are qualitatively similar to those reported in this section.

<sup>15</sup> Since we use analyst-year observations, we cannot control for country and sector fixed effects since many analysts cover stocks across countries and/or sectors.

<sup>16</sup> We obtain similar results when we use the number of countries followed by an analyst rather than a dummy variable to indicate multi-country versus single-country coverage as the dependent variable in our estimated regressions. We do not report these results in the paper for purposes of brevity.

the table are also identical to each other again with the exception of the variable that represents the country factor. They differ from the first two models in one important respect—the independent variables in these regressions represent averages for an analyst’s entire portfolio for a given period. For each of the four models, we also report the change in implied probability of following a multi-country portfolio as an independent variable in the logistic regression changes from quartile 1 (25<sup>th</sup> percentile value) to quartile 4 (75<sup>th</sup> percentile value), holding all other variables constant at their mean values.

Across all four models, the coefficients associated with the country factor are negative and significant at the one percent level, indicating that a higher country factor decreases the likelihood of multi-country coverage. Since a higher value for the country factor implies greater commonalities across firms in the same country, this evidence is consistent with the information efficiency hypothesis. The economic significance of this variable can be gauged by examining the reported change in implied probability that is associated with the parameter estimates. We find that as the country factor rises from its 25<sup>th</sup> percentile value to its 75<sup>th</sup> percentile value, the probability of country specialization increases between 0.036 (3.6%) and 0.119 (11.9%).

The revenue generation hypothesis—specialization is more likely if a market is relatively large—is supported by our results as the parameter estimate associated with the natural logarithm of the country market capitalization variable (*LCTYMV*) is negative and significant in all four models. In fact, *LCTYMV* appears to have the second highest economic significance among all the variables. Specifically, increasing *LCTYMV* from its 25<sup>th</sup> percentile value to its 75<sup>th</sup> percentile value decreases (increases) the probability of country diversification (specialization) by an amount ranging from 0.084 (8.4%) to 0.180 (18.0%).

The parameter estimates associated with broker origin (*BRORIGIN*) are positive and significant. This evidence suggests that employment by brokers domiciled in the U.S. increases

the likelihood of multi-country coverage. A change in the value of the *BRORIGIN* dummy from zero to one increases the probability of multi-country coverage by approximately 11%, suggesting that the economic significance of this variable is relatively high. We also find that likelihood of country specialization is negatively related to broker size (*BRSIZE*). This result runs counter to our predictions regarding this variable. A possible explanation for this result is that larger brokerages try to appeal to a variety of clients. Some of these clients may have an organizational structure that is more compatible with the utilization of multi-country sector-oriented research making it more likely that analysts will cover multiple countries if employed by a large broker (see, for example, Rubino (2003)). Consistent with the analyst experience hypothesis, we find that more experienced analysts are significantly more likely to diversify their coverage across countries. An increase in analyst experience from quartile one to quartile four diminishes the probability of country specialization by approximately 0.03 (3.0%), across all four models.

With the exception of the proportion of idiosyncratic variation in stock returns (*VARRESTOT*) and the extent of accounting disclosure (*ACCTG INDX*), the coefficients on the control variables are consistent across all models and are significant at the one percent level. Recall that we use the natural logarithm of firm equity market capitalization (*LFIRMMV*) and capital market development (*CTYMGDP*) as control variables that capture different aspects of the availability of information. The coefficients associated with both these variables are positive, implying that increased informational availability raises the likelihood of multi-country coverage. Among these variables, firm size has the greatest economic significance as evidenced by the fact that an increase in *LFIRMMV* from quartile one to quartile four increases the probability of multiple country coverage by an amount ranging from 0.098 (9.8%) to 0.133 (13.3%). These results parallel those of Petersen and Rajan (2002) who show that the

geographical distance between banks and their clients changes in response to banks' information processing capabilities. In a similar vein, we find that larger firms and firms located in countries with better disclosure environments are more likely to be followed by multi-country analysts. These results are consistent with the notion that improved information availability, because of better capital market development or larger firm size, reduces the need for geographical proximity between analysts and the stocks they cover.

Our measure of minority shareholder protection, *ANTIDIR RIGHTS* has a negative coefficient and an increase in this variable from its 25<sup>th</sup> percentile value to its 75<sup>th</sup> percentile value reduces the probability of multiple country coverage from 0.127 (12.7%) to 0.176 (17.6%). Better minority shareholder protection may result in higher commission income for the brokerage due to greater interest shown by longer-term investors, i.e., investors who are more likely to value analyst research, thereby making smaller research portfolios more economically viable and specialization more likely.

### *III. B. Analysis of Single versus Multiple Sector Analysts*

Table V presents tests of our hypotheses regarding analyst specialization by sector.<sup>17</sup> The first two models reported in the table are identical with the exception of the variable that represents the sector factor--*VARSECTOT* in the first model and *VARSECSYS* in the second model. All the independent variables employed in these two regressions are average values of the independent variables listed in the first column across all stocks in the analyst's core country-sector for a given period. The third and fourth models reported in the table are also identical to each other again with the exception of the variable that represents the sector factor.

---

<sup>17</sup> We obtain similar results when we use the number of sectors followed by an analyst rather than a dummy variable to indicate multi-sector versus single-sector coverage as the dependent variable in our estimated regressions. We do not report these results in the paper for purposes of brevity.



They differ from the first two models in one important respect—the independent variables in these regressions represent averages for all stocks in an analyst’s portfolio for a given period. To study the economic significance of our results, we also provide information on changes in implied probability. Overall, the economic significance of explanatory variables in our sector regressions is relatively low compared with their power in the country regressions.

Consistent with the information efficiency hypothesis, the parameter estimates associated with the sector factor are negative in all four models. However, they are only significant at conventional levels in Models 2, 3 and, 4. These results are consistent with the notion that a higher sector factor decreases the likelihood of multi-sector coverage. We find that increases in the sector factor from its 25<sup>th</sup> percentile value to its 75<sup>th</sup> percentile value only decreases (increases) the probability of sector diversification (specialization) by between 0.008 (0.8%) and 0.050 (5.0%). The relatively low variation of the sector factor documented in Table III may explain these relatively small changes in implied probabilities.

The revenue generation hypothesis is supported by our results as the parameter estimate associated with the natural logarithm of country-sector market capitalization variable (*LCTYSECMV*) is negative and significant in all four models. Increases in *LCTYSECMV* from quartile one to quartile four raises the probability of sector specialization by an amount ranging from 0.045 (4.5%) to 0.081 (8.1%).

The parameter estimates associated with *BRORIGIN* are negative and significant. The change in implied probability suggests that employment by brokers domiciled in the U.S. increases the likelihood of single-sector coverage by approximately 0.15 (15.0%). We also find that likelihood of sector specialization is positively related to broker size, with an increase in *BRSIZE* from the median value of quartile one to median of quartile four boosting the probability of sector specialization by approximately 0.06 (6.0%). These results are consistent

with our broker culture hypothesis. We find that more experienced analysts are significantly less likely to be specialized, and an increase in analyst experience from its 25<sup>th</sup> percentile value to its 75<sup>th</sup> percentile value diminishes the probability of sector specialization by approximately 0.075 (7.5%) across all four models. This result is consistent with the analyst experience hypothesis.

All the control variables are significant at conventional levels. The coefficient associated with *VARRESTOT* is negative and significant in all four models. This result is consistent with the argument that greater idiosyncratic risk forces an analyst to spend more time and effort to effectively cover a stock, thereby leading to more focused coverage. The likelihood of multiple sector coverage is negatively associated with the natural logarithm of equity market capitalization of the firm (*LFIRMMV*) and accounting index (*ACCTG INDX*). Once again, firm size seems to have the greatest explanatory power as an increase in *LFIRMMV* from quartile one to quartile four reduces the probability of multiple sector coverage by an amount ranging from 0.138 (13.8%) to 0.142 (14.2%). We believe that these relations are consistent with the existence of gains to sector specialization with improved disclosure. For example, improved disclosure allows for analysts to make more meaningful intra-industry comparisons because they have access to more than just basic financial statement information. We also find that specialization by sector is more likely for countries with a higher score for *ANTIDIR RIGHTS*, i.e., countries that offer greater protection for minority shareholders, consistent with our earlier argument that increased shareholder protection will encourage analyst specialization. Finally, greater capital market development is conducive to multi-sector coverage.

In summary, the information efficiency, revenue generation, broker culture, and analyst experience hypotheses all explain, in varying degrees, the choice of analysts/brokerages to follow stocks across countries and sectors. In addition, the inclusion of information on non-core

stocks does not provide any more insights regarding portfolio choice than just an examination of how core characteristics impact the portfolio diversification decision.

#### IV. Forecast Accuracy and Portfolio Choice

In this section, we investigate whether the structure of analysts' research portfolios affects their earnings forecast accuracy. Once again, because we employ analyst-year observations in this analysis, an analyst may appear in our sample in more than one year. We, therefore, employ standard errors that account for the possibility that observations for an analyst across years may be correlated in the analysis that follows (see, for example, Gleason and Lee (2003)).<sup>18</sup> Prior to addressing the above issue, we will first present summary statistics for variables, other than those related to portfolio choice since they were presented earlier in Table III, which the extant literature has argued have an impact on forecast accuracy.

##### *IV.A. Description of Forecast Accuracy and Control Variables*

As is the case in Clement and Tse (2005), we transform all the variables to conduct our examination of the relationship between forecast accuracy and portfolio choice. We undertake these transformations for two reasons. First, the transformed variables control for firm effects by ensuring that every variable is a relative measure for a given stock. Second, since every transformed variable takes on a value between 0 and 1, it ensures that the coefficients in our estimated regressions are comparable across variables.

With the exception of our measures of firm size and forecast accuracy, each raw observation,  $x_{ijt}$ , for analyst  $i$  covering stock  $j$  at time  $t$ , is transformed into a "relative" value  $rx_{ijt}$ , as follows:

---

<sup>18</sup> In additional robustness tests, we also employ standard errors that account for the possibility that the forecasts for analysts employed by the same brokerage are correlated. The results are qualitatively similar.

$$rx_{ijt} = \frac{x_{ijt} - \text{Min}\{x_{jt}\}}{\text{Max}\{x_{jt}\} - \text{Min}\{x_{jt}\}}$$

where  $\text{Min}\{x_{jt}\}$  and  $\text{Max}\{x_{jt}\}$  represent the minimum and maximum values for the variable  $x$  for stock  $j$  during fiscal year  $t$ . These “relative” values are then averaged across stocks in the analyst’s portfolio.

Our measure of analyst forecast accuracy,  $\text{ACCURACY}_{ijt}$  represents the absolute error of analyst  $i$ ’s forecast of stock  $j$ ’s earnings in year  $t$ . Consistent with the literature, only the last forecast issued for the current fiscal year by an analyst at least one month prior to the fiscal year end is employed. To ensure that increasing accuracy of forecasts results in a higher value for the relative variable, the relative value of  $\text{ACCURACY}_{ijt}$  ( $rx_{ijt}$ ), is constructed from each raw observation,  $x_{ijt}$ , using the following procedure:

$$rx_{ijt} = \frac{\text{Max}\{x_{jt}\} - x_{ijt}}{\text{Max}\{x_{jt}\} - \text{Min}\{x_{jt}\}}$$

Once again, these “relative” values are then averaged across stocks in the analyst’s portfolio.

To ensure that we can construct these relative variables, we only consider firms that are covered by at least two analysts. Note that this relative value of the forecast accuracy captures the analyst’s relative forecast accuracy for a given firm. Note also that normalizing the absolute forecast error by either the actual earnings for the fiscal year or by a constant price will not affect the computed values for the scaled accuracy variable.

Summary statistics for the variables employed to model forecast accuracy are presented in Table VI. The statistics presented pertain to the “raw” values of the variables so as to give the reader a better feel for the sample characteristics. Note that all values reported in the table represent averages of the respective raw variables across all stocks in an analyst’s portfolio for a given fiscal year. The variable  $\text{DAYSELPSED}_{ijt}$  measures the length of time in days between

the last earnings forecast by any analyst of stock  $j$ 's fiscal year  $t$  earnings and analyst  $i$ 's forecast of fiscal year earnings. The mean (median) value for  $DAYSELAPSED_{it}$  when averaged across all stocks in an analyst's portfolio is 8.47 (6.00) days. Statistics for  $FORHOR_{it}$ , which measures the average number of days from the date on which analyst  $i$  forecast fiscal year  $t$  earnings for stocks in her portfolio and the last day of fiscal year  $t$ , indicate that the average forecast is issued between three and four months prior to the fiscal year end (mean 119.45 days and median 107.60 days), and that seventy five percent of the forecasts are issued no more than five months prior to the fiscal year end. A mean of 2.26 and a median of 2.00 for  $FORFREQ_{it}$ , which capture the average number of times analyst  $i$  issues forecasts for portfolio stocks during fiscal year  $t$ , suggest that analysts tend to issue two forecasts per year for the stocks they cover, with more active coverage (above the 75<sup>th</sup> percentile) resulting in three or more forecasts per year for a stock.

The next two variables describe characteristics of the analyst. The mean (median) value for  $GENEXP_{it}$  is 841.05 (606.00) indicating that analysts in *IBES International* have been following stocks and reporting to *IBES* for approximately two years. The mean (median) number of stocks covered by an analyst,  $COMPANIES_{it}$  is 8.41 (6.00). The last two variables employed in our analysis describe characteristics of the stocks covered by the analysts and the brokerages employing the analysts. The mean (median) value of \$1,221.70 million (\$1,163.28 million) for  $FIRMMV_{it}$ , the average U.S. dollar value of the equity market capitalization of stocks in analyst  $i$ 's portfolio in year  $t$ , suggests that most of the stocks in our sample have relatively large market capitalizations.<sup>19</sup> This is not surprising given that our sample includes only stocks that are

---

<sup>19</sup> We include  $FIRMMV$  as a control in our tests to control for systematic differences across firms in information availability, analyst incentives, and the level of competition between analysts, as these factors could induce systematic variation in the distributions of the forecast accuracy of analysts across firms.

actively covered by analysts in *IBES*. Given that the 75<sup>th</sup> percentile of the variable is \$3,379.92 million, a comparison with the mean value for the variable suggests that *FIRMMV* is relatively skewed. To account for this skewness, we employ the natural logarithm of *FIRMMV* in our subsequent analysis. *BRSIZE<sub>it</sub>* is a count of the number of analysts employed by analyst *i*'s employing broker in year *t*. On average, brokers employ 155.93 analysts with the smaller brokerages (25<sup>th</sup> percentile or lower) employing under 21 analysts and the larger brokerages employing over 279 analysts (75<sup>th</sup> percentile or higher).

#### *IV.B. Impact of Portfolio Choice on Forecast Accuracy*

Table VII presents regression results that examine the relation between analyst forecast accuracy and portfolio diversity along the country dimension. Model 1 establishes a benchmark against which we compare our results on the relationship between forecast accuracy and portfolio diversity across countries when we control for the endogeneity of portfolio choice. In estimating Model 1, we treat analyst organization as an exogenous variable. The statistically insignificant coefficients associated with the variable *DUMCTY*, a dummy variable that takes the value of 1 if an analyst covers stocks from multiple countries and 0 otherwise, in Model 1 suggests that diversification of coverage across countries has no impact on forecast accuracy. This result supports the notion that the benefits of information complementarities realized offset the loss of scale economies when analysts diversify their coverage across countries.

Models 2 through 5 are estimated under the assumption that country diversification is endogenously determined. We account for the country diversification choice by employing the maximum-likelihood estimation procedure to implement a two-stage Heckman treatment effect methodology. Here the binary self-selection first stage equation is modeled as a probit. Each of these models employs one of the models presented in Table IV as the basis for the country

diversification decision. Model 2 in Table VII employs Model 1 from Table IV to model country diversification. Similarly, Model 3 employs Model 2 from Table IV, Model 4 employs Model 3 from Table IV, and Model 5 employs Model 4 from Table IV. The positive and statistically significant coefficient associated with the variable *DUMCTY* in Models 2 through 5 supports the notion that multi-country research portfolios improve forecast accuracy once the country diversification choice is treated as endogenous.

In all the estimated Heckman models, the coefficient associated with the Inverse Mills Ratio is significantly negative at the 1% level. This negative coefficient implies that brokerages assign multi-country research portfolios to analysts who, for a reason that is not captured by the variables in our forecast accuracy model, are not as likely to forecast accurately.<sup>20</sup> Our results, therefore, indicate that the performance of these analysts is boosted because they are assigned multiple-country portfolios. Alternatively, these results imply that the performance of these analysts would have been worse if they were assigned single-country research portfolios.<sup>21</sup>

The coefficients associated with the control variables in this regression are consistent across all the regressions presented in the table. Further, these coefficients are generally consistent with the evidence presented in prior investigations into the forecast accuracy of analysts. The coefficient estimate associated with *RDAYSELAPSED* indicates that forecasts that

---

<sup>20</sup> For example, as we argued earlier, it is possible that brokers prefer to assign diversified portfolios to analysts who are also responsible for supervising other analysts because the supervisors can better supervise analysts covering different sorts of firms if they themselves cover a variety of firms. Further, these administrative duties may adversely affect the forecasting abilities of the supervisors. An alternative explanation for our result is that supervisors, knowing that their performance may suffer because of their supervisory responsibilities, pick portfolios that help limit the drop off in their performance by selecting performance boosting complex portfolios when they are possible to put together.

<sup>21</sup> Our findings are analogous to finding students who take SAT preparation classes are inherently less capable than students who do not take these classes, but that, by taking the classes, their SAT scores are boosted. Consequently, by not controlling for this self-selection bias, we may find that SAT preparation classes have a non-positive impact on SAT scores.

are clustered together tend to be more accurate. This result is consistent with the findings of Clement and Tse (2005). The variable *RFORHOR* has a significantly negative coefficient, consistent with prior research showing that earnings forecasts closer to the fiscal year end are more accurate (see, e.g., O'Brien, 1988). This result supports the idea that forecasts made later in the fiscal year benefit from the availability of more information and, as a result, tend to be more accurate. The positive coefficient associated with *RFORFREQ*, a proxy for analyst effort, suggests that analysts who work harder produce more accurate results. Jacobs, Lys, and Neale (1999) and Clement (1999) document similar relationships between forecast accuracy with forecast horizon and forecast frequency.

Analyst experience displays a tendency to detract from forecast accuracy as evidenced by the negative and statistically significant coefficient for *RGENEXP*.<sup>22</sup> Our results are generally consistent with prior evidence on the relationship between forecast accuracy and analyst experience. For example, Hong, Kubik, and Solomon (2000) find that forecast accuracy deteriorates with analyst general experience. We had earlier documented that analysts with greater experience are assigned diversified portfolios. Here we find that greater general experience reduces forecast accuracy. Thus, it appears that analysts with greater experience do not have higher ability. We speculate that the responsibilities of these analysts may have evolved over time in that generating more accurate forecast may be just one small part of their job description. It is possible that these analysts are now being evaluated more for their ability to generate revenues for their brokerages and/or have added supervisory responsibilities. Hong, Kubik, and Solomon (2000) find that more experienced analysts are less likely to face

---

<sup>22</sup> Note that the variable *RGENEXP* is different from the variable *GENEXP* that is employed in the first stage of the Heckman procedure and thus its coefficient can be interpreted without accounting for the value of the coefficient of *GENEXP* in the choice model.



unfavorable job outcomes for poorer forecast accuracy than less experience analysts. This result is consistent with the interpretation that the market for analysts recognizes these other responsibilities and, as a result, does not penalize experienced analysts as much for poor performance. In turn, these analysts respond to the incentives provided by the market for analysts.

Our results with respect to *RBRSIZE* suggest that the likely increased availability of resources for analysts working for larger brokers does not translate into more accurate forecasts.<sup>23</sup> We also find that following a larger number of stocks enhances analyst forecast accuracy. Finally, our results suggest that forecasts for larger firms tend to be more accurate. There are a number of potential explanations for this phenomenon, not least of which are the likelihood of improved informational availability for larger firms, greater competition among analysts covering larger firms, and larger rewards for analysts who forecast well for larger firms.

Table VIII presents results from regressions estimated to study the impact of coverage diversification across sectors on forecast accuracy. In Model 1, we examine whether there is any difference in forecast accuracy between analysts who focus their coverage entirely on one sector and analysts who cover stocks from multiple sectors. In this case, analyst organization is treated as an exogenous variable. Once again, the statistically insignificant coefficients associated with the variable, *DUMSEC*, a dummy variable that takes the value of 1 if an analyst covers stocks from multiple sectors and 0 otherwise, suggests that diversification of coverage across sectors has no impact on forecast accuracy. This result supports the notion that the benefits of

---

<sup>23</sup> This result differs from that reported in Clement (1999) who employs a dichotomous classification of brokerages. In his paper, in contrast to our use of a firm level ranking of brokerages, broker size is captured by a dummy variable that takes the value one when the broker is in the highest decile in a global ranking of brokers based on the number of analysts employed in that year. When we employ a broker size classification similar to that in Clement (1999), we also find that the relation between this broker size dummy and forecast accuracy is significantly positive, supporting the hypothesis that the larger brokerages are better able to support analysts. The remaining results remain unchanged.

information complementarities realized offset the loss of scale economies when analysts diversify their coverage across sectors.

Models 2 through 5 also examine the impact of sector diversification on forecast accuracy. Unlike Model 1, however, they are estimated under the assumption that sector diversification is endogenously determined. We account for the sector diversification choice by employing a two-stage Heckman methodology. Each of these models employs one of the models presented in Table V as the model for the sector diversification decision. Model 2 in Table VIII employs Model 1 from Table V to model sector diversification. Similarly, Model 3 employs Model 2 from Table V, Model 4 employs Model 3 from Table V, and Model 5 employs Model 4 from Table V. The positive and statistically significant coefficient associated with the variable *DUMSEC* in Models 2 through 5 supports the notion that multi-sector research portfolios provide more accurate forecasts than single-sector research portfolios once the sector diversification choice is treated as an endogenous variable. Here too the coefficients associated with the Inverse Mills Ratios are negative in all estimated Heckman models but significantly negative at conventional levels in Models 2, 4, and 5. The negative coefficient on the Inverse Mills Ratio implies that analysts who follow multi-sector research portfolios are those who are not likely to forecast accurately. Taken together, these results imply that if multi-sector analysts are assigned to follow single-sector research portfolios, their performance would suffer. Note, however, that results are weaker than those in Table VII. The coefficients associated with the control variables in this regression are consistent across all the regressions presented in the table and are similar to those in Table VII.

Overall, our results indicate that portfolio diversity in the country or sector dimension enhances forecast accuracy once we account for the fact that portfolio choice is endogenously determined. These findings are consistent with the notion that brokerages rationally assign

analysts to follow diversified portfolios as long as the benefits arising from informational complementarities dominate the costs that result from any decrease in economies of scale in information production. The natural question that arises is then why all research portfolios are not diversified either along country and sector lines. The answer may simply lie in the fact the opportunities to exploit informational complementarities may be limited, and where they exist, brokerages organize analyst research to take advantage of them. These opportunities will be greater when the common variation in returns, our proxy for the presence of scale economies, is relatively low for stocks belonging to the same country (sector).

## V. Extensions

In this section we describe results from additional tests designed to provide further insights into the relationship between analyst organization and forecast accuracy. First, we discuss results from a switching regression method to assess the impact of multi-country and multi-sector portfolios on forecast accuracy. The next two sets of tests described in this section are designed to assess the impact of country (sector) diversification while controlling for the sectoral (country) structure of analyst portfolios. All the robustness tests support our earlier conclusions. Regardless of the nature of the test employed, switching to multi-country coverage appears to boost forecast accuracy. The evidence on the positive effects of a switch to multi-sector coverage is somewhat weaker but most of the tests described below suggest that diversifying coverage across sectors tends to boost performance. For purposes of brevity, the results from these tests are not reported in tables in the paper.

### *V.A. Switching Regressions*

In earlier sections, the tests employed to assess the impact of analyst organization on forecast accuracy are valid under the assumption that analyst organization only affects the intercept terms in the regression and that the sensitivity of forecast accuracy to other factors such as forecast horizon and experience is independent of analyst organization. However, it is possible that this assumption is inappropriate. If so, the effect of analyst organization on forecast accuracy can be assessed using a switching regression technique. The only difference between the switching regression technique and the methodology employed thus far is that multiple regressions, one for each form of analyst organization, are estimated to assess the relationship between organization and forecast accuracy. More specifically, as in the tests presented earlier, we first model the choice of analyst portfolios using probit regressions. These

estimates are then employed to estimate the Inverse Mills ratio (*IMR*) for each observation. Finally, in the second stage of the estimation, a separate OLS regression is estimated for each type of analyst organization, i.e., for each form of analyst organization,  $o$ , we estimate the following regression:

$$ACCURACY_o = \beta_{o1} + \beta_{o2} X_2 + \beta_{o3} X_3 + \dots + \beta_{on} X_n + \gamma_o IMR + \varepsilon$$

The effect of switching from analyst organization type  $a$  to analyst organization type  $b$  is then assessed by computing the *Average Treatment Effect (ATE)*, which ignores the effect of the Inverse Mills Ratio and measures the difference in the predicted accuracy across the entire sample of the two forms of analyst organization, i.e.,

$$ATE_{ab} = E\{(\beta_{a1} - \beta_{b1}) + (\beta_{a2} - \beta_{b2}) X_2 + (\beta_{a3} - \beta_{b3}) X_3 + \dots + (\beta_{an} - \beta_{bn}) X_n\}$$

Note that, because our focus is on the effect of portfolio structure on forecast accuracy of the average analyst, we use the *Average Treatment Effect* and not the *Treatment Effect for the Treated* (see, e.g., Hamilton and Nickerson (2003), Heckman and Robb (1986), Rubin (1974), and Rubin (1978) for a distinction between these two effects). Estimating the statistical significance of the average treatment effect using parametric methods is not feasible because we need an estimate from the accuracy regressions of the covariance between the error terms for each type of analyst organization ( $\sigma_{12}$ ) to compute its standard error. The parameter  $\sigma_{12}$  is, however, not estimable in the above kinds of switching regression models (see, for example, Maddala, 1983). We, therefore, resort to non-parametric methods to compute the significance of the average treatment effects.

All results from these switching regression tests of our hypotheses are similar to those reported earlier—complex portfolios tend to boost forecast accuracy, whether the comparison is made between the effect of diversification across sectors or diversification across countries. For example, the average treatment effect is 0.0259 for country diversification when we use the

independent variables in Model 4 of Table IV to model the choice of analyst portfolio along the country dimension. We find that 96.7% of the treatment effects evaluated across all analysts in the sample are non-negative. This proportion is significantly different from 50.0% at the 1% level of significance. Similarly, the average treatment effect for sector diversification is 0.0270 when we use the independent variables in Model 4 of Table V to model the choice of analyst portfolio along the sector dimension. Here, we find that 98.8% of the treatment effects evaluated across all analysts in the sample are non-negative, and that this proportion is significantly different from 50.0% at the 1% level of significance.

#### *V.B. Binomial Tests Employing Finer Classification of Organization Structure*

The tests described earlier compare the impact on forecast accuracy of switching from single country (sector) to multi-country (sector) coverage. Note, however, that some of the analysts who cover stocks from a single country (sector) may cover stocks from multiple sectors (countries) while analysts who cover stocks from multiple countries (sectors) may only focus on a single sector (country). To the extent that changes in sectoral/country coverage can affect performance and there is a systematic relationship between these coverage decisions, tests that assess the impact of country (sector) diversification while controlling for the sector (country) structure of analysts portfolios may result in more accurate conclusions regarding the effect of organization structure on forecast accuracy.

This prompted us to run two sets of robustness tests. First, we employed the Heckman methodology described earlier on appropriate subsets of our sample to estimate the effect of switching from (i) single-country single-sector (SCSS) coverage to single-country multi-sector (SCMS) coverage, (ii) single-country single-sector (SCSS) coverage to multi-country single-sector (MCSS) coverage, and (iii) single-country single-sector (SCSS) coverage to multi-country

multi-sector (*MCMS*) coverage. The results from these tests support our earlier conclusions that diversifying coverage across countries and sectors results in improved forecast accuracy.

*V.C. Multinomial Tests Employing Finer Classification of Organization Structure*

In the final set of tests, we extend the switching regression methodology described earlier to examine the effect on forecast accuracy of the following four forms of analyst organization: *SCSS*, *SCMS*, *MCSS*, and *MCMS*. To estimate the average treatment effects of these four organizational forms, we first use the variables employed earlier to estimate multinomial logistic regressions that model analyst choices among these four organizational forms with *SCSS* coverage as the base group. Coefficient estimates associated with the variables measuring commonalities within countries and sectors, country and sector sizes, and broker origin conform closely both in terms of sign and statistical significance with our earlier estimates of coefficients associated with these variables. The four regressions estimating forecast accuracy employ the Inverse Mills Ratios generated from the multinomial regression to control for the endogeneity in analyst organization.<sup>24</sup> In all cases, estimates of the average treatment effects obtained from pairs of these forecast accuracy regressions (*SCSS* vs. *SCMS*, *SCSS* vs. *MCSS*, and *SCSS* vs. *MCMS*) suggest that diversification across countries helps improve forecast accuracy. The evidence on the positive effects of diversification across sectors is somewhat weaker but generally suggests that diversification across sectors also helps improve forecast accuracy.

---

<sup>24</sup> Lee (1982) demonstrates how non-normal unobservables implied by the multinomial regressions can be transformed into normal variables, thus, easily allowing us to construct the inverse Mills ratios. His approach provides us with a simple method to account for multinomial choices.

## VI. Conclusions

A widely held belief among financial economists is that analysts focus their research coverage on individual sectors. For example, Ross, Westerfield, and Jaffe (2005, p. 337) contend that security analysts, "...are employed by brokerage houses to follow the companies in individual industries. For example, an analyst for a particular brokerage house might follow all the firms in, say, the auto industry." While this sectoral approach may largely reflect the manner in which analyst research is organized in the U.S., we find rich cross-sectional variation in analyst research portfolios along both country and sector dimensions for analysts covered by International *IBES*. In fact, this variation reflects a long-standing debate among practitioners on whether international analyst specialization should occur along country or sector dimensions (see Rudd (1989)). We shed light on this debate by conducting an in-depth study of the economic forces and constraints that determine the structure of the research portfolios of international analysts.

We develop four non-mutually exclusive hypotheses relating to analyst research portfolios. First, the *information efficiency hypothesis* posits that the strength of commonalities between firms within a country or sector determine the nature of specialization. Second, the *revenue generation hypothesis* posits that the choice between focused versus broad coverage of stocks by analysts is determined by opportunities to generate revenue streams for the brokerage. Next, the *broker culture hypothesis* predicts that the observed cross-sectional variation in analyst specialization reflects broker culture in terms of how brokerages have traditionally organized their research in their home economies. Finally, the *analyst experience hypothesis* states that more experienced analysts, either due to following different types of stocks over the course of their careers or because of their greater ability, will tend to cover more diverse portfolios.



Our empirical tests provide broad support for all these hypotheses. Specifically, consistent with the *information efficiency hypothesis*, we find that the probability that analysts specialize by country is higher as the country factor rises, and the likelihood that they focus on a sector increases as the sector factor strengthens. In line with the *revenue generation hypothesis*, we find that higher country equity market capitalization increases the likelihood of country specialization; while a larger total equity market capitalization of a sector in a particular country (country-sector market capitalization) increases the likelihood that analysts will cover firms in a single sector. In support of the *broker culture hypothesis*, we document that analysts employed by U.S. brokers are less likely to specialize by country and more likely to specialize by sector. Finally, consistent with the *analyst experience hypothesis*, we find that more experienced analysts have less focused research portfolios.

When we examine the relation between forecast accuracy and portfolio organization using a two-stage Heckman methodology to control for the endogeneity in the structure of analyst research portfolios, we find that diversified portfolios promote forecast accuracy. This is the case when we consider portfolio diversification across either countries or sectors. In contrast, Clement (1999) documents a negative relation between forecast accuracy and measures of portfolio complexity for a sample of U.S. analysts and Clement, Rees, and Swanson (2003) find no relation between forecast accuracy and measures of portfolio complexity for a sample of international analysts. In both these studies, the structures of analyst portfolios are treated as being exogenously determined. Thus, our analysis indicates that a simple comparison of forecast accuracy across analysts with different portfolio structures without controlling for self-selection in portfolio choice appears to understate differences in forecast accuracy. Our finding that diversified research portfolios outperform focused research portfolios is consistent with the notion that brokerages rationally assign analysts to follow diversified portfolios as long as the

marginal benefits arising from informational complementarities dominate the marginal costs that result from any decrease in economies of scale in information production. We, therefore, conclude that brokerages structure analyst research portfolios in a manner that is conducive to enhancing forecast accuracy.

Our results provide a fresh perspective on the scandals that have linked various conflicts of interest faced by analysts to the quality of their research. The general impression is that brokerage research and recommendations are entirely driven by investment-banking related incentives and have little relation to fundamental values. We weigh in on this issue not only by examining how the production of research is structured but by also studying the impact of portfolio choice on brokerage output. We do find that the potential for investment banking business and/or brokerage commissions impacts the manner in which analyst research is organized. However, we also find that the structure of the production process of research in brokerages is consistent with producing good quality research.

## REFERENCES

- Bacmann, J., and G. Bolliger, 2001, Who are the best? Local versus foreign analysts on the Latin American stock markets, University of Neuchatel and FAME working paper.
- Bae, K.H., R. Stulz, and H. Tan, 2005, Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts, Ohio State University and Queen's University working paper.
- Berger, A.N, N.H. Miller, M.A. Petersen, R.G. Rajan, and J.C. Stein, 2005, Does function follow organizational form? Evidence from the lending practices of large and small firms, *Journal of Financial Economics* 76, 237-269.
- Bolliger, G., 2001, The characteristics of individual analysts' forecasts in Europe, FAME working paper.
- Boni, L., and K.L. Womack, 2005, Analysts, industries, and price momentum, University of New Mexico working paper.
- Chang, C., 2003, Information footholds: Expatriate analysts in an emerging market, University of California at Berkeley working paper.
- Choe, H., C. Kho, and R.M. Stulz, 2000, Do domestic investors have more valuable information about individual stocks than foreign investors? *Review of Financial Studies* 18, 795-829.
- Clement, M.B., 1999, Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27, 285-303.
- Clement, M.B., L. Rees, and E.P. Swanson, 2003, The influence of culture and corporate governance on the characteristics that distinguish superior analysts, *Journal of Accounting, Auditing, and Finance* 18, 593-618.
- Clement, M.B., and S.Y. Tse, 2005, Financial analyst characteristics and herding behavior in forecasting, *Journal of Finance* 60, 307-341.
- Chen, Z., 2003, Buy-side and sell-side: The industrial organization of information production in securities industry, Wharton School working paper.
- Coval, J.D., and T.J. Moskowitz, 2001, The geography of investment: Informed trading and asset prices, *Journal of Political Economy* 109, 811-841.
- DeAngelo, L., 1981, Auditor size and Auditor quality, *Journal of Accounting and Economics* 3, 183-199.
- Garicano, L., and T. N. Hubbard, 2005, Specialization, firms and markets: The division of labor within and between law firms, University of Chicago working paper.

- Gilson, S.C., P.M. Healy, C.F. Noe, and Krishna G. Palepu (2001), Analyst specialization and conglomerate stock breakups, *Journal of Accounting Research* 39, 565-582.
- Gleason, C., and C. Lee, 2003, Analyst forecast revisions and market price discovery, *Accounting Review* 78, 193-225.
- Greene, W.H., 2000, *Econometric Analysis* 5<sup>th</sup> Edition, NJ: Prentice Hall, Inc.
- Griffin, J.M., and A.G. Karolyi, 1998, Another look at the role of the industrial structure of markets for international diversification strategies, *Journal of Financial Economics* 50, 351-373.
- Grinblatt, M., and M. Keloharju, 2000, The investment behavior and performance of various types of investors: A study of Finland's unique data set, *Journal of Financial Economics* 55, 43-67.
- Hamilton, B.H., and J.A. Nickerson, 2003, Correcting for endogeneity in strategic management research, *Strategic Organization* 1, 51-78.
- Heckman, J., 1978, Dummy endogenous variables in a simultaneous equations system, *Econometrica* 46, 931-961.
- Heckman, J., 1979, Sample selection bias as a specification error, *Econometrica* 47, 153-161.
- Heckman, J., and R. Robb, 1986, Alternative methods for solving the problem of selection bias in evaluating the impact of treatments on outcomes, In *Drawing inferences from self-selected samples*, edited by H. Wainer, 63-107, Mahwah, NJ: Lawrence Erlbaum Associates.
- Heston, H.L., and G.K. Rouwenhorst, 1994, Does industrial structure explain the benefits of international diversification? *Journal of Financial Economics* 36, 3-27.
- Hong, H., and J.D. Kubik, 2003, Analyzing the analyst: Career concerns and biased forecasts, *Journal of Finance* 58, 313-351.
- Hong, H., J.D. Kubik, and A. Solomon, 2000, Security analysts' career concerns and herding of earnings forecasts, *RAND Journal of Economics* 31, 121-144.
- Huber, P., 1967, The behavior of maximum likelihood estimators under non standard conditions, In *Procedures of the Fifth Annual Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, edited by L.M. LeCam and J. Neyman, 221-233, Berkeley, CA: University of California Press.
- International Accounting and Auditing Trends*, Princeton, NJ: Center for International Financial Analysis and Research, 1995.
- Jacob, T., T. Lys, and M. Neale, 1999, Expertise in forecasting performance of security analysts, *Journal of Accounting and Economics* 28, 51-82.

- Kang, J., and R.M. Stulz, 1997, Why is there a home bias? An analysis of foreign portfolio equity ownership in Japan, *Journal of Financial Economics* 46, 2-28.
- Kennedy, P., 1986, Interpreting dummy variables, *Review of Economics and Statistics* 68, 174-175.
- Kini, O, S.L. Mian, M.J. Rebello, and A. Venkateswaran, 2006, Core specialization, forecast accuracy, and the market for international analysts, Georgia State University working paper.
- La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R.W. Vishny, 1997, Legal determinants of external finance, *Journal of Finance* 52, 1131-1150.
- Lee, L.F., 1982, Some approaches to the correction of selectivity bias, *Review of Economic Studies* 49, 355-372.
- Li, K., and N. Prabhala, 2005, Self-selection models in corporate finance, University of Maryland working paper.
- Maddala, G.S., 1983, *Limited-dependent and qualitative variables in econometrics*, Cambridge: Cambridge University Press.
- Malloy, C.J., 2005, The geography of equity analysis, *Journal of Finance* 60, 719-755.
- Morck, R., B. Yeung, and W. Yu, 2000, The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics* 58, 215-260.
- O'Brien, P., 1988, Analysts' forecasts as earnings expectations, *Journal of Accounting and Economics* 10, 53-85.
- Petersen, M.A., and R.G. Rajan, 2002, Does distance still matter? The information revolution in small business lending, *Journal of Finance* 57, 2533-2570.
- Rajan, R.G., and L. Zingales, 1998, Financial dependence and growth, *American Economic Review* 88, 559-586.
- Reuters Institutional Investor Survey*, Institutional Investor Research Group, 2002.
- Rogers, W.H., 1993, Regression standard errors in clustered samples, *Stata Technical Bulletin* 13, 19-23.
- Ross, S.A., R.W. Westerfield, and J. Jaffe, *Corporate Finance* 7<sup>th</sup> edition, New York, NY: McGraw-Hill Irwin, 2005.
- Rubin, D., 1974, Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of Educational Psychology* 66, 688-701.

- Rubin, D., 1978, Bayesian inference for causal effects, *The Annals of Statistics* 6, 34-58.
- Rubino, J., 2003, "The new global industry analysis," *CFA Magazine*, July-August, 46-47.
- Rudd, A., 1989, Industry versus country correlation, *Journal of Portfolio Management* 15, 79-88.
- Seasholes, M., 2004, Re-examining information asymmetries in emerging markets, University of California at Berkeley working paper.
- Sonney, F., 2005, Sector versus country specialization and financial analysts' performance, University of Neuchatel and FAME working paper.
- Stigler, G., 1951, Division of labor is limited by the extent of the market, *Journal of Political Economy* 59, 185-193.
- White, H, 1980, A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity, *Econometrica* 48, 817-838.
- Wooldridge, J.M., 2002, *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.
- World Stock Exchange Fact Book*, Morris Plains, NJ: Electronic Commerce, Inc., 1996-2003.
- Zhou, J. and Ronald Elder, 2003, Audit firm size, industry specialization, and earnings management by IPO firms, State University of New York at Binghamton and Syracuse University working paper.

**Table I**  
**Determinants of Analyst Research Portfolio Structures**

This table presents the predicted relations between the probability of country/sector diversification and various independent variables under our four hypotheses. The variables employed to explain analyst diversification by country (sector) include the following variables: *VARCTYTOT* (*VARSECTOT*) the proportion of total weekly return variability for each firm that is attributable to the movements in the weekly returns on the country (sector) factor. *VARCTYSYS* (*VARSECSYS*) is the ratio of the variance of the weekly returns on the country (sector) factor to the sum of the variances of the weekly returns on the global, country, and sector factors. *CTYMV* (*CTYSECMV*) is the stock market capitalization of the country (country-sector) of the firm. *BRSIZE* is measured as the number of analysts employed by the broker. *BRORIGIN* is a dummy variable that takes the value of one if the broker is of U.S. origin, and is zero otherwise. *ANALEXP* is the number of days elapsed between the first ever forecast made by an analyst and her latest forecast for that particular fiscal year on International *IBES*.

Hypothesis	Proxy Variable	Hypothesized sign	
		Country diversification	Sector diversification
Information efficiency hypothesis			
	Country factor ( <i>VARCTYTOT</i> and <i>VARCTYSYS</i> )	(-)	n.a.
	Sector factor ( <i>VARSECTOT</i> and <i>VARSECSYS</i> )	n.a.	(-)
Revenue generation hypothesis			
	Country market capitalization ( <i>CTYMV</i> )	(-)	n.a.
	Country-sector market capitalization ( <i>CTYSECMV</i> )	n.a.	(-)
Broker culture hypothesis			
	Broker origin ( <i>BRORIGIN</i> )	(+)	(-)
	Broker size ( <i>BRSIZE</i> )	(-)	(-)
Analyst experience hypothesis			
	Number of days on <i>IBES</i> ( <i>GENEXP</i> )	(+)	(+)

**Table II**

This table presents mean and median values of characteristics of analyst research portfolios for a sample of 19,379 analysts covering 11,292 firms between 1996 and 2002. The characteristics described include the number of countries in which firms followed by an analyst are domiciled, the number of sectors in which firms followed by an analyst operate based on the *SIG* codes on International *IBES*, the number of firms followed, and analyst experience. This data is presented in three panels. Panel A contains a description of analyst portfolios for our entire sample of analysts. Panel B presents this information for analysts whose entire research portfolio consists of firms domiciled in a single country (*SC*) and those covering firms in multiple countries (*MC*). In Panel C descriptions are for single-sector analysts (*SS*) and multiple sector analysts (*MS*). We report t-statistics (Wilcoxon z-statistics) for differences between mean (median) characteristics across analyst groups in each panel. <sup>a</sup>, <sup>b</sup>, and <sup>c</sup> indicate significance at the 0.01 0.05, and 0.10 levels, respectively.

Analyst group	Number of analyst years	Percentage of total	Number of countries followed by analyst		Number of sectors followed by analyst		Number of firms followed by analyst		Analyst experience (days)	
			Mean	Median	Mean	Median	Mean	Median	Mean	Median
Panel A: All										
<i>ALL</i>	47,049	100	1.46	1.00	2.36	2.00	8.41	6.00	841	606
Panel B: By country										
<i>SC</i>	36,697	78.00	1.00	1.00	2.40	2.00	8.27	6.00	814	569
<i>MC</i>	10,352	22.00	3.07	3.00	2.22	2.00	8.91	8.00	934	726
t-stat (z-stat)			350.00 <sup>a</sup>	214.92 <sup>a</sup>	-11.09 <sup>a</sup>	-4.99 <sup>a</sup>	9.71 <sup>a</sup>	24.93 <sup>a</sup>	14.53 <sup>a</sup>	18.66 <sup>a</sup>
Panel C: By sector										
<i>SS</i>	18,714	39.78	1.43	1.00	1.00	1.00	5.47	3.00	781	496
<i>MS</i>	28,335	60.22	1.47	1.00	3.26	3.00	10.35	9.00	880	672
t-stat (z-stat)			4.22 <sup>a</sup>	5.11 <sup>a</sup>	280.00 <sup>a</sup>	191.84 <sup>a</sup>	94.82 <sup>a</sup>	100.37 <sup>a</sup>	13.98 <sup>a</sup>	25.58 <sup>a</sup>



Table III

### Descriptive Statistics on Variables that are Hypothesized to Impact Portfolio Structure

The sample consists of 47,049 analyst-year observations over the period 1996-2002. Each of the explanatory variables described below is measured in two ways. First, their values are averaged only for the stocks in the core country-sector in the analyst's portfolio (Core selection). Second, their values are averaged across all stocks in the analyst's portfolio (Overall selection). The variables employed to explain analyst specialization include the proportion of total weekly return variability for each firm that is attributable to the movements in the weekly returns on the country factor (*VARCTYTOT*), the ratio of the variance of the weekly returns on the country factor to the sum of the variances of the weekly returns on the global, country, and sector factors (*VARCTYSYS*), the proportion of total weekly return variability for each firm that is attributable to the movements in the weekly returns on the sector factor (*VARSECTOT*), and the ratio of the variance of the weekly returns on the sector factor to the sum of the variances of the weekly returns on the global, country, and sector factors (*VARSECSYS*). *CTYMV* is the stock market capitalization of the country in which the firm is domiciled. *CTYSECMV* is the country-sector market capitalization for the stock. *BRORIGIN* is a dummy variable that takes the value of one if the broker is of U.S. origin, and is zero otherwise. *BRSIZE* is measured as the number of analysts employed by the broker. *GENEXP* is the number of days elapsed between the first ever forecast made by an analyst and his/her latest forecast during a given year on International *IBES*. *VARRESTOT* is the residual variance of the firm divided by its total variance, where both are estimated using weekly returns. *FIRMMV* represents the firm equity market capitalization. Following LLSV (1997), we also employ proxies for the levels of information disclosure (*ACCTG INDX*), minority shareholder protection (*ANTIDIR RIGHTS*), and capital market development (*CTYMVGDP*).

Variable	Core Selection		Overall Selection	
	Mean	Median	Mean	Median
<i>VARCTYTOT</i>	0.199	0.148	0.201	0.156
<i>VARCTYSYS</i>	0.553	0.540	0.557	0.543
<i>VARSECTOT</i>	0.066	0.043	0.066	0.049
<i>VARSECSYS</i>	0.320	0.195	0.324	0.242
<i>CTYMV</i> (\$ m.)	458,630.429	530,194.925	437,136.158	479,260.706
<i>CTYSECMV</i> (\$ m.)	51,534.151	62,255.365	46,536.862	53,369.783
<i>BRORIGIN</i> (proportion)	0.243	0.000	0.243	0.000
<i>BRSIZE</i> (# of analysts)	155.925	107.000	155.925	107.000
<i>GENEXP</i> (days)	841.053	606.000	841.053	606.000
<i>FIRMMV</i> (\$ m.)	1,208.337	1,152.859	1,221.702	1,163.281
<i>ACCTG INDX</i>	68.773	69.000	68.548	69.000
<i>ANTIDIR RIGHTS</i>	3.407	4.000	3.372	4.000
<i>CTYMVGDP</i>	0.839	0.429	0.803	0.451

Table IV

**Binomial Logistic Regression Analysis for Single Country versus Multiple Country Analysts**

The dependent variable is a dummy variable, *DUMCTY*, that takes the value of one if the analyst is a multiple country analyst, and is zero otherwise. Our unit of analysis is the analyst-year combination. Under the core selection models, all the independent variables for a year are computed by averaging the value of the underlying variable across all stocks in an analyst's core country-sector, while under the overall selection models, all the independent variables for a year are computed by averaging the value of the underlying variable across all stocks in an analyst's portfolio. The variables employed to explain analyst specialization by country include the proportion of total weekly return variability for each firm that is attributable to the movements in the weekly returns on the country factor (*VARCTYTOT*) and the ratio of the variance of the weekly returns on the country factor to the sum of the variances of the weekly returns on the global, country, and sector factors (*VARCTYSYS*). *LCTYMV* is the natural logarithm of stock market capitalization of the country the firm in which the firm is domiciled. *BRORIGIN* is a dummy variable that takes the value of one if the broker is of U.S. origin, and is zero otherwise. *BRSIZE* is measured as the number of analysts employed by the broker. *GENEXP* is the number of years elapsed between the first ever forecast made by an analyst and his/her latest forecast during our 1995-1996 period on International *IBES*. *VARRESTOT* is the residual variance of the firm divided by its total variance, where both are estimated using weekly returns. *LFIRMMV* represents the natural logarithm of firm equity market capitalization. Following LLSV (1997), we also employ proxies for the levels of information disclosure (*ACCTG INDX*), minority shareholder protection (*ANTIDIR RIGHTS*) and capital market development (*CTYMGDP*). The change in implied probability is computed by varying the independent variable in the logistic regression from quartile 1 (25<sup>th</sup> percentile value) to quartile 4 (75<sup>th</sup> percentile value), holding all other variables constant at their mean values. The p-values are in parentheses and reflect standard errors that are corrected for the cluster sample problem arising from multiple observations across years for each analyst. <sup>a</sup>, <sup>b</sup>, and <sup>c</sup> indicate significance at the 0.01, 0.05, and 0.10 levels, respectively.

Variable	Core Selection				Overall Selection			
	Model 1 (N = 44,671)		Model 2 (N = 44,671)		Model 3 (N = 45,984)		Model 4 (N = 45,984)	
	Coeff.	Implied Prob.	Coeff.	Implied Prob.	Coeff.	Implied Prob.	Coeff.	Implied Prob.
Intercept	0.615 <sup>a</sup> (0.01)	n.a.	4.080 <sup>a</sup> (0.00)	n.a.	2.407 <sup>a</sup> (0.00)	n.a.	5.909 <sup>a</sup> (0.00)	n.a.
<i>Year Dummies</i>	Yes	n.a.	Yes	n.a.	Yes	n.a.	Yes	n.a.
<i>VARCTYTOT</i>	-1.328 <sup>a</sup> (0.00)	-0.036			-1.794 <sup>a</sup> (0.00)	-0.052		
<i>VARCTYSYS</i>			-2.036 <sup>a</sup> (0.00)	-0.115			-2.264 <sup>a</sup> (0.00)	-0.119
<i>LCTYMV</i>	-0.330 <sup>a</sup> (0.00)	-0.084	-0.504 <sup>a</sup> (0.00)	-0.134	-0.524 <sup>a</sup> (0.00)	-0.141	-0.704 <sup>a</sup> (0.00)	-0.180
<i>BRORIGIN</i>	0.661 <sup>a</sup> (0.00)	0.100	0.694 <sup>a</sup> (0.00)	0.110	0.688 <sup>a</sup> (0.00)	0.111	0.724 <sup>a</sup> (0.00)	0.110
<i>BRSIZE</i>	0.001 <sup>a</sup> (0.00)	0.035	0.002 <sup>a</sup> (0.00)	0.072	0.002 <sup>a</sup> (0.00)	0.074	0.002 <sup>a</sup> (0.00)	0.069
<i>GENEXP</i> (x 10 <sup>2</sup> )	0.024 <sup>a</sup> (0.00)	0.029	0.022 <sup>a</sup> (0.00)	0.028	0.025 <sup>a</sup> (0.00)	0.032	0.024 <sup>a</sup> (0.00)	0.029
<i>VARRESTOT</i>	-0.068 (0.11)	-0.003	0.050 (0.24)	0.002	-0.104 <sup>c</sup> (0.08)	-0.005	0.083 <sup>c</sup> (0.07)	0.004
<i>LFIRMMV</i>	0.343 <sup>a</sup> (0.00)	0.098	0.330 <sup>a</sup> (0.00)	0.099	0.449 <sup>a</sup> (0.00)	0.133	0.431 <sup>a</sup> (0.00)	0.119
<i>ACCTG INDX</i>	0.000 (0.96)	0.000	-0.006 (0.39)	-0.009	0.001 (0.68)	0.002	-0.004 (0.27)	-0.006
<i>ANTIDIR RIGHTS</i>	-0.440 <sup>a</sup> (0.00)	-0.176	-0.357 <sup>a</sup> (0.00)	-0.149	-0.484 <sup>a</sup> (0.00)	-0.167	-0.398 <sup>a</sup> (0.00)	-0.127
<i>CTYMGDP</i>	0.535 <sup>a</sup> (0.00)	0.050	0.557 <sup>a</sup> (0.00)	0.055	0.563 <sup>a</sup> (0.00)	0.057	0.585 <sup>a</sup> (0.00)	0.055
$\chi^2$	3,219.84 <sup>a</sup> (0.00)		3,370.16 <sup>a</sup> (0.00)		3,748.28 <sup>a</sup> (0.00)		3,908.84 <sup>a</sup> (0.00)	
Pseudo R <sup>2</sup> (%)	14.63		15.75		17.21		18.37	

Table V

### Binomial Logistic Regression Analysis for Single Sector versus Multiple Sector Analysts

The dependent variable is a dummy variable, *DUMSEC*, that takes the value of one if the analyst is a multiple sector analyst, and is zero otherwise. Our unit of analysis is the analyst-year combination. Under the core selection models, all the independent variables for a year are computed by averaging the value of the underlying variable across all stocks in an analyst's core country-sector, while under the overall selection models, all the independent variables for a year are computed by averaging the value of the underlying variable across all stocks in an analyst's portfolio. The variables employed to explain analyst specialization by sector include the proportion of total weekly return variability for each firm that is attributable to the movements in the weekly returns on the sector factor (*VARSECTOT*) and the ratio of the variance of the weekly returns on the sector factor to the sum of the variances of the weekly returns on the global, country, and sector factors (*VARSECSYS*). *LCTYSECMV* is the stock market capitalization of the sector for the country in which the firm is domiciled. *BRORIGIN* is a dummy variable that takes the value of one if the broker is of European origin, and is zero otherwise. *BRSIZE* is measured as the number of analysts employed by the broker. *GENEXP* is the number of years elapsed between the first ever forecast made by an analyst and his/her latest forecast during our 1995-1996 period on International IBES. *VARRESTOT* is the residual variance of the firm divided by its total variance, where both are estimated using weekly returns. *LFIRMMV* represents the natural logarithm of the firm equity market capitalization. Following LLSV (1997), we also employ proxies for the levels of information disclosure (*ACCTG INDX*), minority shareholder protection (*ANTIDIR RIGHTS*) and capital market development (*CTYMV GDP*). The change in implied probability is computed by varying the independent variable in the logistic regression from quartile 1 (25th percentile value) to quartile 4 (75th percentile value), holding all other variables constant at their mean values. The p-values are in parentheses and reflect standard errors that are corrected for the cluster sample problem arising from multiple observations across years for each analyst. <sup>a</sup>, <sup>b</sup>, and <sup>c</sup> indicate significance at the 0.01, 0.05, and 0.10 levels, respectively.

Variable	Core Selection				Overall Selection			
	Model 1 (N = 44,671)		Model 2 (N = 44,671)		Model 3 (N = 45,984)		Model 4 (N = 45,984)	
	Coeff.	Implied Prob	Coeff.	Implied Prob	Coeff.	Implied Prob	Coeff.	Implied Prob
Intercept	4.167 <sup>a</sup> (0.00)	n.a.	4.465 <sup>a</sup> (0.00)	n.a.	4.809 <sup>a</sup> (0.00)	n.a.	5.074 <sup>a</sup> (0.00)	n.a.
Year Dummies	Yes	n.a.	Yes	n.a.	Yes	n.a.	Yes	n.a.
<i>VARSECTOT</i>	-0.425 (0.15)	-0.008			-0.778 <sup>b</sup> (0.02)	-0.013		
<i>VARSECSYS</i>			-0.659 <sup>a</sup> (0.00)	-0.050			-0.698 <sup>a</sup> (0.00)	-0.049
<i>LCTYSECMV</i>	-0.084 <sup>a</sup> (0.00)	-0.045	-0.107 <sup>a</sup> (0.00)	-0.057	-0.139 <sup>a</sup> (0.00)	-0.070	-0.159 <sup>a</sup> (0.00)	-0.081
<i>BRORIGIN</i>	-0.622 <sup>a</sup> (0.00)	-0.153	-0.610 <sup>a</sup> (0.00)	-0.150	-0.626 <sup>a</sup> (0.00)	-0.153	-0.617 <sup>a</sup> (0.00)	-0.151
<i>BRSIZE</i>	-0.001 <sup>a</sup> (0.00)	-0.062	-0.001 <sup>a</sup> (0.00)	-0.062	-0.001 <sup>a</sup> (0.00)	-0.061	-0.001 <sup>a</sup> (0.00)	-0.062
<i>GEXP</i> (x 10 <sup>2</sup> )	0.034 <sup>a</sup> (0.00)	0.076	0.035 <sup>a</sup> (0.00)	0.077	0.034 <sup>a</sup> (0.00)	0.075	0.034 <sup>a</sup> (0.00)	0.075
<i>VARRESTOT</i>	-0.143 <sup>a</sup> (0.00)	-0.011	-0.166 <sup>a</sup> (0.00)	-0.013	-0.162 <sup>a</sup> (0.00)	-0.012	-0.153 <sup>b</sup> (0.04)	-0.012
<i>LFIRMMV</i>	-0.278 <sup>a</sup> (0.00)	-0.142	-0.273 <sup>a</sup> (0.00)	-0.139	-0.285 <sup>a</sup> (0.00)	-0.138	-0.283 <sup>a</sup> (0.00)	-0.139
<i>ACCTG INDX</i>	-0.006 <sup>a</sup> (0.00)	-0.016	-0.006 (0.00)	-0.016	-0.005 <sup>b</sup> (0.02)	-0.013	-0.006 <sup>b</sup> (0.00)	-0.016
<i>ANTIDIR RIGHTS</i>	-0.135 <sup>a</sup> (0.00)	-0.098	-0.124 <sup>a</sup> (0.00)	-0.090	-0.136 <sup>a</sup> (0.00)	-0.071	-0.126 <sup>a</sup> (0.01)	-0.066
<i>CTYMV GDP</i>	0.103 <sup>a</sup> (0.00)	0.020	0.109 <sup>a</sup> (0.00)	0.021	0.107 <sup>a</sup> (0.00)	0.021	0.109 <sup>a</sup> (0.00)	0.021
$\chi^2$	1,878.03 <sup>a</sup> (0.00)		1,976.75 <sup>a</sup> (0.00)		1,987.29 <sup>a</sup> (0.00)		2,074.84 <sup>a</sup> (0.00)	
Pseudo-R2 (%)	6.74		7.08		7.20		7.50	

**Table VI**  
**Descriptive Statistics on Raw Control Variables in Accuracy Regressions**

The sample consists of 47,049 analyst-year observations over the period 1996-2002. All the control variables are computed for each year by averaging the value of the underlying variable across all stocks in an analyst's portfolio. *DAYSELAPSED* measures the length of time in days between the last earnings forecast by any analyst of firm *j*'s fiscal year *t* earnings and analyst *i*'s forecast of fiscal year earnings. *FORHOR* measures the number of days from the date on which analyst *i*'s forecast fiscal year *t* earnings for firm *j* and the last day of fiscal year *t*. *FORFREQ* is a proxy for the intensity with which an analyst covers a firm. It is estimated by the number of times analyst *i* issues forecasts for firm *j* during fiscal year *t*. *GENEXP* measures the number of days between an analyst's first forecast in the *IBES* database and her last forecast for year *t*. *BRSIZE* represents the number of analysts working for analyst *i*'s employer during fiscal year *t* in which she issues a forecast of firm *j*'s earnings. *COMPANIES* is a count of the number of stocks in analyst *i*'s research portfolio in year *t*. The final control variable *FIRMMV* is the equity market capitalization of firm *j* during fiscal year *t* during which it is covered by analyst *i*.

Variable	25 <sup>th</sup> Percentile	Mean	Median	75 <sup>th</sup> Percentile
<i>DAYSELAPSED</i> (days)	2.667	8.472	6.000	11.500
<i>FORHOR</i> (days)	75.667	119.453	107.600	149.600
<i>FORFREQ</i>	1.333	2.264	2.000	3.000
<i>GENEXP</i> (days)	271.000	841.053	606.000	1,184.000
<i>BRSIZE</i> (# of analysts)	21.000	155.925	107.000	279.000
<i>COMPANIES</i> (# of stocks)	3.000	8.409	6.000	11.000
<i>FIRMMV</i> (\$ m.)	436.217	1,221.702	1,163.281	3,379.929

**Table VII**  
**Forecast Accuracy for Multi- versus Single-Country Analysts**

This table presents fixed effects regression models with average *relative* forecast accuracy (*ACCURACY*) of analyst *i* for fiscal year *t* as the dependent variable for a sample of analyst-year observations made over the period 1996-2002. The average scaled accuracy is computed across all the stocks in the analyst's portfolio. Model 1 estimates an ordinary least squares regression. Models 2 through 5 estimate the same relations but employ the Heckman's two-stage procedure to control for the endogeneity in analyst portfolio choice (multi- versus single-country analyst) in the first stage. These models differ only in the first-stage selection model employed. In Models 2 and 3, the first-stage selection models employ independent variables that are averaged across all stocks in an analyst's core with *VARCTYTOT* (*VARCTYSYS*) as the proxy for the country factor in Model 2 (3). In Models 4 and 5, the first-stage selection models employ independent variables that are averaged across all stocks in an analyst's portfolio with *VARCTYTOT* (*VARCTYSYS*) as the proxy for the country factor in Model 4 (5). In all these models, we control for year fixed effects. *DUMCTY* is a dummy variable that takes the value one if the analyst follows stocks in more than one country; otherwise it takes the value zero. *DAYSELAPSED* measures the length of time in days between the last earnings forecast by any analyst of firm *j*'s fiscal year *t* earnings and analyst *i*'s forecast of fiscal year earnings. *RFORHOR* measures the number of days from the date on which analyst *i*'s forecast fiscal year *t* earnings for firm *j* and the last day of fiscal year *t*. *RFORFREQ* is a proxy for the intensity with which an analyst covers a firm. It is estimated by the number of times analyst *i* issues forecasts for firm *j* during fiscal year *t*. *GENEXP* measures the number of days between an analyst's first forecast in the *IBES* database and her last forecast for year *t*. *BRSIZE* represents the number of analysts working for analyst *i*'s employer during fiscal year *t* in which she issues a forecast of firm *j*'s earnings. *COMPANIES* is a count of the number of stocks in analyst *i*'s research portfolio in year *t*. The final control variable *LFIRMMV* is the natural logarithm of average equity market capitalization of firm *j* during fiscal year *t* during which it is covered by analyst *i*. All control variables are *relative* with the exception of *LFIRMMV* and, as such, have the prefix "R" preceding their names. The p-values are in parentheses and reflect standard errors that are corrected for the cluster sample problem arising from multiple observations across years for each analyst. Subscripts have been dropped in the table below for ease of presentation. <sup>a</sup>, <sup>b</sup>, and <sup>c</sup> indicate significance at the 0.01, 0.05, and 0.10 levels, respectively.

	Core Selection			Overall Selection	
	Model 1	Model 2	Model 3	Model 4	Model 5
	OLS (N = 45,947)	Heckman ( <i>VARCTYTOT</i> )	Heckman ( <i>VARCTYSYS</i> )	Heckman ( <i>VARCTYTOT</i> )	Heckman ( <i>VARCTYSYS</i> )
Intercept	0.601 <sup>a</sup> (0.00)	0.624 <sup>a</sup> (0.00)	0.621 <sup>a</sup> (0.00)	0.634 <sup>a</sup> (0.00)	0.630 <sup>a</sup> (0.00)
<i>Year Dummies</i>	Yes	Yes	Yes	Yes	Yes
<i>DUMCTY</i>	-0.003 (0.26)	0.068 <sup>a</sup> (0.00)	0.060 <sup>a</sup> (0.00)	0.081 <sup>a</sup> (0.00)	0.070 <sup>a</sup> (0.00)
<i>RDAYSELAPSED</i>	-0.070 <sup>a</sup> (0.00)	-0.078 <sup>a</sup> (0.00)	-0.077 <sup>a</sup> (0.00)	-0.073 <sup>a</sup> (0.00)	-0.073 <sup>a</sup> (0.00)
<i>RFORHOR</i>	-0.231 <sup>a</sup> (0.00)	-0.234 <sup>a</sup> (0.00)	-0.234 <sup>a</sup> (0.00)	-0.234 <sup>a</sup> (0.00)	-0.234 <sup>a</sup> (0.00)
<i>RFORFREQ</i>	0.039 <sup>a</sup> (0.00)	0.040 <sup>a</sup> (0.00)	0.039 <sup>a</sup> (0.00)	0.041 <sup>a</sup> (0.00)	0.041 <sup>a</sup> (0.00)
<i>RGENEXP</i>	-0.015 <sup>a</sup> (0.00)	-0.022 <sup>a</sup> (0.00)	-0.021 <sup>a</sup> (0.00)	-0.023 <sup>a</sup> (0.00)	-0.021 <sup>a</sup> (0.00)
<i>RBRSIZE</i>	-0.009 <sup>a</sup> (0.00)	-0.016 <sup>a</sup> (0.00)	-0.015 <sup>a</sup> (0.00)	-0.017 <sup>a</sup> (0.00)	-0.016 <sup>a</sup> (0.00)
<i>RCOMPANIES</i>	0.006 (0.19)	0.012 <sup>a</sup> (0.00)	0.012 <sup>a</sup> (0.00)	0.013 <sup>a</sup> (0.00)	0.012 <sup>a</sup> (0.00)
<i>LFIRMMV</i>	0.015 <sup>a</sup> (0.00)	0.011 <sup>a</sup> (0.00)	0.011 <sup>a</sup> (0.00)	0.009 <sup>a</sup> (0.00)	0.010 <sup>a</sup> (0.00)
Adj. R <sup>2</sup> (%)	9.20				
$\chi^2$		3,143.14 <sup>a</sup> (0.00)	3,144.43 <sup>a</sup> (0.00)	3,224.78 <sup>a</sup> (0.00)	3,212.20 <sup>a</sup> (0.00)
Inverse Mills Ratio		-0.044 <sup>a</sup> (0.00)	-0.040 <sup>a</sup> (0.00)	-0.054 <sup>a</sup> (0.00)	-0.048 <sup>a</sup> (0.00)

**Table VIII**  
**Forecast Accuracy for Multi- versus Single-Sector Analysts**

This table presents fixed effects regression models with average *relative* forecast accuracy ( $ACCURACY_{it}$ ) of analyst  $i$  for fiscal year  $t$  as the dependent variable for a sample of analyst-year observations made over the period 1996-2002. Model 1 estimates an ordinary least squares regression. Models 2 through 5 estimate the same relations but employ the Heckman's two-stage procedure to control for the endogeneity in analyst portfolio choice (multi- versus single-sector analyst) in the first stage. These models differ only in the first-stage selection model employed. In Models 2 and 3, the first-stage selection models employ independent variables that are averaged across all stocks in an analyst's core with  $VARSECTOT$  ( $VARSECSYS$ ) as the proxy for the country factor in Model 2 (3). In Models 4 and 5, the first-stage selection models employ independent variables that are averaged across all stocks in an analyst's portfolio with  $VARSECTOT$  ( $VARSECSYS$ ) as the proxy for the country factor in Model 4 (5).  $DUMSEC$  is a dummy variable that takes the value one if the analyst follows stocks in more than one country; otherwise it takes the value zero.  $DAYSELAPSED$  measures the length of time in days between the last earnings forecast by any analyst of firm  $j$ 's fiscal year  $t$  earnings and analyst  $i$ 's forecast of fiscal year earnings.  $FORHOR$  measures the number of days from the date on which analyst  $i$ 's forecast fiscal year  $t$  earnings for firm  $j$  and the last day of fiscal year  $t$ .  $FORFREQ$  is a proxy for the intensity with which an analyst covers a firm. It is estimated by the number of times analyst  $i$  issues forecasts for firm  $j$  during fiscal year  $t$ .  $GENEXP$  measures the number of days between an analyst's first forecast in the *IBES* database and her last forecast for year  $t$ .  $RBSIZE$  represents the number of analysts working for analyst  $i$ 's employer during fiscal year  $t$  in which she issues a forecast of firm  $j$ 's earnings.  $COMPANIES$  is a count of the number of stocks in analyst  $i$ 's research portfolio in year  $t$ . The final control variable  $LFIRMMV$  is the natural logarithm of average equity market capitalization of firm  $j$  during fiscal year  $t$  during which it is covered by analyst  $i$ . All control variables are *relative* with the exception of  $LFIRMMV$  and, as such, have the prefix "R" preceding their names. The p-values are in parentheses and reflect standard errors that are corrected for the cluster sample problem arising from multiple observations across years for each analyst. Subscripts have been dropped in the table below for ease of presentation. <sup>a</sup>, <sup>b</sup>, and <sup>c</sup> indicate significance at the 0.01, 0.05, and 0.10 levels, respectively.

	Core Selection			Overall Selection	
	Model 1	Model 2	Model 3	Model 4	Model 5
	OLS (N = 45,947)	Heckman ( $VARSECTOT$ ) (N = 43,794)	Heckman ( $VARSECSYS$ ) (N = 43,794)	Heckman ( $VARSECTOT$ ) (N = 44,999)	Heckman ( $VARSECSYS$ ) (N = 44,999)
Intercept	0.600 <sup>a</sup> (0.00)	0.584 <sup>a</sup> (0.00)	0.590 <sup>a</sup> (0.00)	0.575 <sup>a</sup> (0.00)	0.582 <sup>a</sup> (0.00)
<i>Year Dummies</i>	Yes	Yes	Yes	Yes	Yes
<i>DUMSEC</i>	0.002 (0.38)	0.017 <sup>b</sup> (0.02)	0.012 <sup>c</sup> (0.10)	0.024 <sup>a</sup> (0.00)	0.019 <sup>b</sup> (0.01)
<i>RDAYSELAPSED</i>	-0.070 <sup>a</sup> (0.00)	-0.078 <sup>a</sup> (0.00)	-0.078 <sup>a</sup> (0.00)	-0.074 <sup>a</sup> (0.00)	-0.074 <sup>a</sup> (0.00)
<i>RFORHOR</i>	-0.231 <sup>a</sup> (0.00)	-0.232 <sup>a</sup> (0.00)	-0.232 <sup>a</sup> (0.00)	-0.232 <sup>a</sup> (0.00)	-0.232 <sup>a</sup> (0.00)
<i>RFORFREQ</i>	0.039 <sup>a</sup> (0.00)	0.039 <sup>a</sup> (0.00)	0.039 <sup>a</sup> (0.00)	0.040 <sup>a</sup> (0.00)	0.040 <sup>a</sup> (0.00)
<i>RGENEXP</i>	-0.015 <sup>a</sup> (0.00)	-0.019 <sup>a</sup> (0.00)	-0.018 <sup>a</sup> (0.00)	-0.019 <sup>a</sup> (0.00)	-0.018 <sup>a</sup> (0.00)
<i>RBSIZE</i>	-0.009 <sup>a</sup> (0.00)	-0.008 <sup>b</sup> (0.02)	-0.008 <sup>b</sup> (0.01)	-0.007 <sup>b</sup> (0.04)	-0.008 <sup>b</sup> (0.03)
<i>RCOMPANIES</i>	0.004 (0.42)	0.006 (0.16)	0.006 (0.17)	0.005 (0.26)	0.005 (0.27)
<i>LFIRMMV</i>	0.015 <sup>a</sup> (0.00)	0.016 <sup>a</sup> (0.00)	0.015 <sup>a</sup> (0.00)	0.016 <sup>a</sup> (0.00)	0.016 <sup>a</sup> (0.00)
Adj. R <sup>2</sup> (%)	9.20				
$\chi^2$		3,137.44 <sup>a</sup> (0.00)	3,135.25 <sup>a</sup> (0.00)	3,176.67 <sup>a</sup> (0.00)	3,174.19 <sup>a</sup> (0.00)
Inverse Mills Ratio		-0.010 <sup>b</sup> (0.02)	-0.007 (0.15)	-0.014 <sup>a</sup> (0.00)	-0.011 <sup>b</sup> (0.02)

### Appendix A: Extraction of Country and Sector Factors

We followed the methodology described in Heston and Rouwenhorst (1994) to extract country and sector factors. The underlying model for weekly stock returns for security  $i$  that belongs to country  $j$  and sector  $k$  is assumed to be as follows:

$$R_{it} = \alpha_t + \beta_{jt} + \gamma_{kt} + \varepsilon_{it},$$

where  $\alpha_t$  is the base level of return in period  $t$ ,  $\beta_{jt}$  and  $\gamma_{kt}$  are the country and sector effects respectively, and  $\varepsilon_{it}$  is the firm specific component of returns. The firm-specific component of returns has zero mean and finite variance and is uncorrelated across securities.

Our data consists of weekly returns for firms in 45 countries and 11 sectors.<sup>25</sup> For each of the 104 weeks in our sample period, we estimated the following cross-sectional dummy variable regression using weighted least squares:

$$R_{it} = \alpha_t + \beta_{1t}C_{i1} + \beta_{2t}C_{i2} + \dots + \beta_{44t}C_{i44} + \gamma_{1t}S_{i1} + \gamma_{2t}S_{i2} + \dots + \gamma_{10t}S_{i10} + \varepsilon_{it},$$

where  $C_{ij}(S_{ik})$  is dummy variable that takes the value of one if the firm belongs to country  $j$  (sector  $k$ ), and zero otherwise. The weights employed in the regressions are the market percentage capitalizations of the securities at the beginning of each week.

The number of country (sector) dummies in the regression is one less than the number of countries (sectors) in our sample. Following Kennedy (1986) and Heston and Rouwenhorst (1994), the country and sector factors for the remaining country and sector were estimated using the following two restrictions:

---

<sup>25</sup> To decompose returns, we used a subset of the *Datastream Master* file that imposes additional screens to capture either obvious data entry errors or stocks that trade very infrequently. Firms were dropped from the sample if they did not meet the following screens: (i) returns and market capitalization data were available for less than half the sample period and (ii) more than half the available market capitalization and returns data had zero values. The weekly returns data on the resulting sample of firms was used to extract global, country, and sector factors.

$$\sum_{j=1}^{45} w_{jt} \beta_{jt} = 0; \sum_{k=1}^{11} w_{kt} \gamma_{kt} = 0,$$

where  $w_{jt}$  ( $w_{kt}$ ) is the ratio of the market capitalization of all firms in country  $j$  (sector  $k$ ) and the market capitalization of all the firms in the sample in week  $t$ . This procedure enabled us to construct a time-series of 104 weekly returns on global, sector, and country factors.<sup>26</sup>

---

<sup>26</sup> As a check, we replicate Table 3 in Heston and Rouwenhorst (1994) for a sample consisting of 12 European countries and 9 sectors. Specifically, we decompose the variance of the excess return of a country index into a pure country factor and a component that is the variance of the sum of the 9 sector factors. We also decompose the variance of the sector index return into a pure sector effect and sum of the 12 country effects. These decompositions yield results similar to what they report in their paper.