determinants of the quantity and nature of financial data presented at Web sites (see [4, 5]). To the best of our knowledge, however, only one published study [6] has investigated the timeliness, as opposed to the extent, of online financial data. That study employs a sample of 47 firms, and tracks the time elapsed from the dates the firms' Form 10-K financial reports are filed with the SEC until they appear on the firms' Web sites. The study found an average lag of 30 days, with shorter lags associated with: greater corporate profitability, shorter lags in announcing earnings via press releases, and the use of multiple file formats for annual report/Form 10-K presentation. Longer lags are associated with external links to the SEC's Web-based EDGAR archives.

The authors of [6] interpret their evidence as suggesting the following conclusions. Managers of more profitable firms have greater incentives to rapidly disseminate their firms' financial results. Earnings release lags proxy for incentives to rapidly disseminate financial information not captured by other model variables. Use of multiple formats for 10-K/annual report files characterizes managers committed to providing high-quality IR materials, including rapid updating of financial reports online. Managers view external links to their firms' 10-Ks at the EDGAR site as substitutes for providing the 10-Ks at their firms' own Web sites. Finally, the authors of [6] argue that the capability to rapidly disseminate financial information online is irrelevant if managers do not have adequate incentives for rapid dissemination. They view the 30-day average update lag as indicating low incentives to update.

This study extends prior studies in a variety of ways. Prior research studied factors affecting the updating of one IR item at corporate Web sites: the 10-K/annual report. We investigate the updating of all IR items at corporate Web sites. The dependent variables are summarized in Panel A of Table 1. "Newest Content" equals the number of days between our Web site visit and the most recent dated material in a firm's IR pages. The variable "Oldest Content" equals the number of days between our Web site visit and the oldest dated material in a firm's IR pages. Stale material left at IR Web sites potentially creates legal liability [11]. A third variable, "Average Freshness," was investigated, but models explaining this variable were not statistically significant.

Using Web crawler software allows us to dramatically increase sample size to 421 firms, versus 47 employed by [6]. The software automated data collection, capturing the structure and age profile of the IR portion of the firms' Web sites.

This study adopts an innovative approach to investigating how large corporations manage their investor relations Web site. Rather than studying the impact of the investor relations Web site on the firm's valuation, this study investigates the impact of various characteristics of the firm on the quality of its Web-based disclosures.

Research Sample and Procedures

Companies comprising the 2002 Fortune 500 list were used for this study. Financial information was obtained from Compustat. Using corporate IR addresses as starting points, Web information was collected using WebCat, a Web crawler program [7]. WebCat recursively searches and documents each site's content and structure. Pages containing targeted keywords (see Figure 1) are dubbed "interesting," which triggers the recursive search of all previously nonvisited hyperlinks found on that page. Failure to find any of these terms terminates the search along this branch of the tree.

This study focuses on that portion of the IR site under the direct control of the firm, as indicated by the Web page's hosting domain. Each link is designated as Home-Based if the hosting domain name is associated with the targeted firm; otherwise it is designated as Outsourced.

File freshness is defined as the difference between the file download and upload dates reported in the file header. Some Web technologies made it impossible to determine the upload time of the Web page. After adjusting for time zone changes, 18% of the files were found to have an infeasible freshness value of negative one implying that the file was uploaded after it was downloaded in the study. The cause is likely a synchronization problem between the local computer used in the study and the remote-hosting servers. All lags were incremented by one for the analyses.

Data collection issues with the Web crawler software reduced the sample size to 421 from the original 500 firms. Three firms were eliminated due to mergers. Seventeen firms did not appear to have an IR Web site. The Web technology used on another 59 firms prevented the proper analysis of the site by the Web crawler software. Additionally, missing financial data for five firms further reduced the sample size to 416.

The explanatory variables, summarized in Panel B of Table 1, fall into two categories. First, we employ three variables to proxy for economic incentives prompting rapid IR updates (firm size, firm health, equity issuance). We anticipate that economic incentives are negatively associated with update lags.

Previous research indicates that larger firms are likely to provide more financial information, both in traditional media [3, 9] and online [5]. These findings suggest that the benefits of information dissemination