Data Appendix

Most of the accounting items were taken directly from Datastream, as described in the Data Appendix Table. Adjustments were made to obtain a consistent series for investment, and to estimate the net capital stock at replacement cost.

Our measure of investment records gross fixed investment. For the period up to and including 1991 we use ds435, which was Datastream's total new fixed assets variable, from the sources and uses accounts. This variable is the sum of ds431 (flow of expenditure on gross direct purchases of fixed assets) and ds432 (the book value of gross acquisitions of fixed assets). Following the change in UK accounting procedures in 1991, ds431 is relabelled ds1024 and ds432 is no longer available. In place of ds432 we use $\max[0, \text{ds479}]$, where ds479 is as ds432, except net of the book value of divestments of fixed assets.

Our capital stock variable $p_t^K K_{it}$ is a replacement cost estimate of net fixed assets at current prices, where p_t^K is an implicit price deflator for capital goods. We take the book value of net fixed assets (ds339) in the first sample year for a company, and adjust for pre-sample capital goods price inflation making the assumption that all capital is three years old. The ds339 variable comprises land, buildings, plant and machinery, and excludes inventories. For subsequent years we allow for depreciation and inflation according to the perpetual inventory formula

$$p_{t+1}^K K_{it+1} = \left[(1-\delta)\, p_t^K K_{it}\right] \frac{p_{t+1}^K}{p_t^K} + p_{t+1}^K I_{it+1} \tag{1}$$

with $\delta = 0.08$, where $p_t^K I_{it}$ is gross fixed investment, as described above. I thank Steve Bond for providing the Fortran program that computes this formula.

The investment price series p_t^K is formed using Office for National Statistics (ONS) data as the weighted sum of price series for buildings and for plant/machinery, where the weights reflect the respective shares in aggregate investment.

Our 2-digit SIC output price series, taken from ONS publications, was kindly provided by Daphne Nicolitsas. Firms were allocated to SIC groups based on their Stock Exchange sector classification.

Our initial sample is a panel of 767 nonfinancial companies that were listed on the London Stock Exchange in December 1987, and that did not die before June 1988. Various 'cleaning' procedures were applied to the data. Companies with the following data characteristics were removed entirely from the dataset: companies with fewer than two observations; companies with missing series on key variables (such as sales revenue and employment); and companies with fewer than 50 employees. These procedures reduced the sample from an initial 767 companies to 687 companies. Our rumours data begin in 1988; this further reduces the sample from 687 to 643 companies.

Implausible employment observations were replaced by figures from Extel where possible.[1] We also examined series with large jumps in the key production function variables, namely sales, employment and the capital stock. The standard approach is to exclude companies with large jumps (for example year-on-year changes of a factor exceeding three). Whilst this approach may be sensible for some empirical questions, it does not seem appropriate here because if two companies merge then we would fully expect sales, employment and capital to increase, possibly by a large amount. Thus we looked for jumps in *ratios* between the key variables, namely the ratios of sales-to-labour, capital-to-sales, and capital-to-labour. The logarithms of these ratios are roughly symmetrically distributed, and a jump of a factor of three in the ratio itself corresponds approximately to a change in the logarithm of the ratio of at least five standard deviations from the mean. Sixty-six companies had changes of more than a factor of three in these ratios at some point during the sample period. In each case we assessed whether the jumps were implausible (for example with reference to reports in the *Financial Times*), and if so we excluded the appropriate series.

Observations on accounting year periods of less than 11 months or more than 13 months were removed. Only consecutive yearly observations were kept. If an accounting year in a company's series was either missing or relates to a financial year that is shorter than 11 months or longer than 13 months, then the observation was removed,

[1] We use Datastream rather than Extel as our primary source of accounting data because Extel generally removes dead companies soon after exit, while Datastream provides reasonably full coverage of dead firms (at least for our period of interest, namely 1989-96). For a study of takeovers, this is obviously a decisive factor in Datastream's favour.

and the longer post-1988 series was kept. For example, suppose a company reports in accounting years 1988-95, but that the 1990 observation is either missing or relates to a non-12-month period. Then we removed the 1988-90 observations, but kept the 1991-95 series. Thus although all of our companies were listed on the London Stock Exchange in January 1988, the *data series* for some of our companies do not begin until after 1988. This explains the fact that in Column 1 of Table 1 there are only 579 companies with data for 1988, even though there were 687 companies in our sample.

Details of which firms were the subject of successful takeover bids were obtained from the London Share Price Database, which records company exits, death date, and reason for exits. However, for takeovers the LSPD death date refers to the listing cancellation, which often occurs several months after the actual bid is accepted by the target shareholders. Hence for each takeover exit we checked the *Financial Times* for the correct date of bid acceptance, and used this date as the death date.

The LSPD also provides a Stock Exchange sector classification variable. Scott Evans kindly provided an algorithm that maps this classification into a 2-digit SIC industry code. We used this algorithm to allocate each of our companies to a 2-digit SIC industry.

For the purposes of data analysis, it was necessary to create a 'year' variable for each observation. If a company's financial year ends between January and June, then this observation is classified as belonging to the previous calendar year; if a company's financial year ends between July and December, then this observation is classified as belonging to the current calendar year. For example, consider a company whose financial year ends in January. Then data from company accounts (for example, sales, employment, debt etc) for the 12 month period ending in January 1988 are classified as 1987 data by our 'year' variable. However, for a company whose financial year ends in December, data from company accounts that are reported in December 1988 are classified as 1988 by our 'year' variable.

Information on takeover speculation was taken from the *Financial Times* CD-ROM for 1988-95.[2] For each company in the sample a search was performed for the keywords

[2] The *Financial Times* CD-ROM is not available for years before 1988.

'speculation', 'rumours' or 'bid talk', and was cross-referenced with various merger keywords such as 'takeover', 'bid' and 'acquisition'. Each hit resulting from this search was read, and a unit (date-specific) observation recorded for each article in which a company was touted as a potential takeover target.[3] For the purpose of regression analysis, these date-specific observations were aggregated into a yearly count for each company.[4] We found that grouping all observations where there were 3 or more rumours gave a slightly improved fit compared to using the raw rumour count variable, reflecting the fact that the distribution of the rumour count is very skewed (see Figure 1). Hence in the reported regression analysis we used a specification where values $0, 1, 2, 3$ correspond to 0 rumours, 1 rumour, 2 rumours and 3 or more rumours respectively. However, all our results were robust to alternative specifications of the measure of rumours.

[3] We also distinguished 'serious' from 'less serious' speculation, where serious speculation names a bidder as well as the target; this distinction proved to be uninformative in the regression analysis.

[4] Consistent with the use of company accounts data in the study, this yearly count refers to the accounting year of each company, not the calendar year.