

## Details on the Active Share data file

(c) Antti Petajisto, 2009-01-20

### Variable definitions

|                |   |
|----------------|---|
| wfici          | Wharton financial institution code number; unique fund identifier   |
| rdate          | Report date of fund's holdings  |
| fundno         | Fund number; supposedly unique fund identifier but occasionally reused in later periods                                       |
| icdi           | ICDI number; unique identifier for one share class of a fund; here we include the ICDI of the largest share class of a fund   |
| index_min      | Benchmark index that produces the lowest Active Share for a fund as of each report date                                       |
| activeshare    | Active Share of a fund relative to the benchmark index (index_min) as of each report date                                     |
| trackingerror  | Tracking error volatility of a fund, computed based on the prior 6 months of daily returns; expressed as an annualized number |
| indexfund      | Indicator variable for a pure index fund  |
| enhanced_index | Indicator variable for an enhanced index fund   |
| TNA            | Total net assets of a fund based on its U.S. stock holdings; only includes securities matched with the CRSP stock file        |
| ticker         | Ticker symbol of the largest share class of the fund  |
| fund_name      | Name of the largest share class of the fund   |

### Other details

All variables are defined and the sample of funds is selected as in Cremers and Petajisto (2008). The benchmark indices included in the calculations were 5 indices from S&P, 12 indices from Russell, and 2 indices from DJ Wilshire. For example, "S5" means S&P500, "W4" means Wilshire 4500, and "R2G" means Russell 2000 Growth. This sample extends the original sample by three additional years (2004-2006).

When these numbers were computed, I only had Russell index data until 6/2006 and Wilshire data until 12/2004. As a result, Wilshire and Russell indices drop out after these dates. For example, the Vanguard Total Stock Market Fund has W5 as its benchmark until 12/2004, then R3 from 1/2005 to 6/2006, and S5 from 7/2006 to 12/2006. Of course, this is not a concern for S&P-benchmarked funds, nor for Russell-benchmarked funds until after 6/2006.

The benchmark index is the one that produces the lowest Active Share for a fund at each point in time. This sometimes differs from the fund's official self-declared benchmark index (as stated in its prospectus).