Industry Betas and Equilibrium Models:
An Alternate Approach to Calculating Risk Measures

Andrew MacNamara

The Leonard N. Stern School of Business
Glucksman Institute for Research in Securities Markets
Faculty Advisor: Stephen J. Brown

April 2, 2007

I. Introduction

According to standard financial theory, investors in financial assets are, on average, compensated for the risk of a particular asset by its distribution-inclusive return. Many equilibrium models attempt to capture and quantify this idea, from the Capital Asset Pricing Model (CAPM) to various incarnations of the Arbitrage Pricing Theory (APT).

In the CAPM, risk is measured by beta, a statistical construct designed to capture the amount by which a liquid financial asset's returns change in relation to the returns on "the market." The APT model takes this a step further, relating the return on a stock to a set of "factors" that represent macroeconomic effects, weighted by each stock's exposure to each factor, or *factor loadings*.

Much academic research has focused on testing these equilibrium models in an attempt to determine whether they adequately describe the markets and to determine which model describes markets best. Lintner[1] used regressions to determine betas for a set of 301 stocks from 1954 to 1963. Then he performed a cross-sectional regression to test the security market line, regressing each stock's return over the period against its beta and its residual risk. He found evidence that the residual risk is priced, contradicting the CAPM's predictions. In response, Miller and Scholes[2] critiqued some of the statistical problems with Lintner's model, and found that the misestimation of betas caused a significant problem. Black, Jensen, and Scholes[3] attempted to

---

[1] Douglas, George. *Risk in the Equity Markets: An Empirical Appraisal of Market Efficiency*. (Ann Arbor, Mich.: University Microfilms, Inc., 1968).
[2] Miller, M.H. and M. Scholes. "Rates of Return in Relation to Risk: A Re-Examination of Some Recent Findings," in Jensen, M. (ed.). *Studies in the Theory of Capital Markets* (New York: Praeger, 1972).
[3] Black, F., M.C. Jensen, and M. Scholes. "The Capital Asset Pricing Model: Some Empirical Tests," in Jensen, M. (ed.). *Studies in the Theory of Capital Markets* (New York: Praeger, 1972).

correct this by forming decile portfolios of stocks to reduce this misestimation; they found strong support of the two-factor or zero-beta form of the CAPM. Finally, Fama and MacBeth[4] extended the second pass cross-sectional regression analysis, performing it for each month in their study rather than across the entire time period, and testing other implied hypotheses of the CAPM. Again, their tests provide support for the CAPM.

While most of these tests focused on the U.S. markets, other studies were done internationally. It was found that the traditional equilibrium models did *not* fit the Japanese markets nearly as well as they fit the U.S. markets. In 1990, Brown & Otsuki performed a study of an APT model on the Japanese markets, but with a twist: they allowed the factor loadings for the Japanese stocks to be related to the industries to which the Japanese companies were exposed.[5] Thus, the factor loadings for the stocks, rather than being estimated with a regression and being fixed for each company over the entire period of study, were allowed to vary as the Japanese companies changed their industry exposure through investment and divestment. Brown & Otsuki found that this modified equilibrium model fit the Japanese markets to a degree that was comparable with traditional studies of the U.S. markets.

One presumption of the study is that the high degree to which Japanese companies changed their industry exposures over the period studied caused the traditional equilibrium models to "fail" and the modified model to "work." In other words, it was not the fact that Japanese companies tended to be exposed to more industries than U.S. companies, but rather the fact that they changed those industry exposures so much more than U.S. companies. To see why,

---

[4] Fama, Eugene, and J. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy.* 71 (May/June 1973). pp. 607-636.

[5] Brown, Stephen J. and Toshiyuki Otsuki. "Macroeconomic Factors and the Japanese Equity Markets: The CAPMD Project." *Japanese Capital Markets*. Ballinger Publishing Co. 1990.

consider a conglomerate firm. If it does not change its industry exposures much during a particular time period, a traditional regression of the firm's returns on the market index or on a set of APT factors should capture the effect of diversification for that time period. However, if it changes those industry exposures, a traditional regression would estimate only an average of the effects of its industry exposures over time, and thus would be subject to significant error.

This study relates to other types of research on equilibrium models. For example, many research papers have examined the proposition that companies' risk changes over time. Cho and Engle have shown that CAPM betas vary predictably over time.[6] Blume[7] and Levy[8] have shown that company betas tend to converge on the market beta of one as companies become more mature. This study allows the risk factors to change over time as well, and it might be interesting to revisit these other types of studies to see how much of the time-varying nature of betas is due to changing industry exposure and how much is due to other factors.

Another related idea in finance is that of a "bottom-up" or fundamental beta. According to Aswath Damodaran, a company's beta is related to whether the company's products or services are discretionary or not and the degree of leverage, both operating and financial, with which the company operates.[9] This study aims to show that a company's risk is related to the industries in which it operates, taking care of one of these fundamental factors.

---

[6] Cho, Young-Hye and Robert F. Engle. "Time-Varying Betas and Asymmetric Effects of News: Empirical Analysis of Blue Chip Stocks." Under revision. February 2000.
[7] Blume, Marchall. "Betas and Their Regression Tendencies." *Journal of Finance*. Vol. X, No. 3 (June 1975). pp. 785 – 795.
[8] Levy, Robert. "On the Short-Term Stationarity of Beta Coefficients." *Financial Analysts Journal*. Vol. 27, No. 5 (Dec. 1971). pp. 55 – 62.
[9] Damodaran, Aswath. *Investment Valuation.* 2nd Edition. New York: John Wiley & Sons. 2002. p. 193.

Given the similarity to other concepts in finance, I believe that risk measures that vary with time in relation to a company's industry exposure should, on average, improve the power of the equilibrium model under study to explain the market's returns.

II. Methodology

The traditional way to calculate industry returns is to form portfolios of stocks based on the main reported line of business in which the relevant company operates. However, this methodology ignores the fact that many companies operate in several industries. In the U.S. markets, conglomerates like this have become much more rare, but in the Japanese markets studied by Brown & Otsuki, this was a major problem.

Their solution was to impute the returns on the industries by a statistical method that considers all the public companies in the economy as well as each one's reported industry exposures. These data were not publicly available at the time, but Brown & Otsuki were granted access to a dataset specially compiled to include just this information. Luckily, U.S. public companies are required to report segmentation in their regulatory filings.

I decided that though U.S. companies tend not to be segmented as much as Japanese companies, including the reported segmentation for each company rather than just the main reported line of business is a more "correct" methodology and is more in line with the goals of the study, which aim to determine how the company's risk changes with changes in industry exposure. Therefore, this was the method I used to compute industry returns in my study.

Specifically, I assumed that on average, companies' returns were a linear combination of industry returns, weighted by the companies' exposures to those industries. The variable I used as a proxy for "exposure" was percentage of total sales reported by the company to have come from a particular industry. In other words:

$$R_i = b_{1,i} \times I_1 + b_{2,i} \times I_2 + \ldots + b_{N,i} \times I_N$$, where:

- R_i is the return on a particular stock.
- $b_{j,i}$ is the i^{th} company's exposure to the j^{th} industry.
- I_j is the return on the j^{th} industry.

This looks remarkably like a cross-sectional regression model, where we know the company returns and the company exposures and the regression coefficients would give us the industry returns. However, this model suffers from a problem: all the *b* variables add up to 100%, a violation of the basic assumptions of ordinary least squares regression. Therefore, I transformed the model to correct this.

Since we know the sum of the *b* variables is 100%, we can write $b_{N,i} = 1 - b_{1,i} - b_{2,i} - \ldots - b_{N-1,i}$. Substituting this back into the original model, we obtain:

$$R_i = b_{1,i} \times I_1 + b_{2,i} \times I_2 + \ldots + b_{N-1,i} \times I_{N-1} + \left(1 - b_{1,i} - b_{2,i} - \ldots - b_{N-1,i}\right) \times I_N .$$

Rearranging,

$$R_i = b_{1,i} \times \left(I_1 - I_N\right) + b_{2,i} \times \left(I_2 - I_N\right) + \ldots + b_{N-1,i} \times \left(I_{N-1} - I_N\right) + I_N .$$

This model does not suffer from the linear combination effect. In fact, if we used this model to perform a cross-sectional regression on all stocks for a given time period, with the *b* variables (from 1 to $N - 1$) as independent variables, we would obtain regression coefficients that generally represent the difference between a particular industry's return and the N^{th} industry's return, and a constant term that represented the N^{th} industry's return. Deriving the actual industry returns from these coefficients is simply a matter of adding back the constant term to the regression coefficients corresponding to the independent variables. As described below, this technique is exactly how I derived the industry returns for each period of time.

Additionally, a key assumption of this study relates to the formation of the factor loadings for a particular company based on the factor loadings for a company. Specifically, I assume that the set of factor loadings for a portfolio of securities is equal to a set of weighted averages of factor loadings for the individual portfolio companies, weighted by their portfolio weights. Thus, if we consider a company as a portfolio of industry exposures, we should be able to calculate that company's overall set of factor loadings as a weighted average of the factor loadings of its component industries.

I chose the Fama-French factor model, a form of APT, as the main equilibrium model for this study. This model is specified as:

$R_i = a + b_{RM-RF}\left(R_m - R_f\right) + b_{SMB}\left(SMB\right) + b_{HML}\left(HML\right) + \varepsilon$, where:[10]

- $R_m - R_f$, the excess return on the market, is the value-weighted return on all NYSE, AMEX, and NASDAQ stocks (from CRSP) minus the one-month Treasury bill rate (from Ibbotson Associates).
- SMB (Small Minus Big) is the average return on the three small portfolios minus the average return on the three big portfolios.
- HML (High Minus Low) is the average return on the two value portfolios minus the average return on the two growth portfolios.

The *b* variables in this APT model are factor loadings, and these are the main point of my study. As described below, I calculated two sets of factor loadings for each stock: one set using

[10] From Kenneth French's Data Library page: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html

the traditional regression method and the other set using another method that accounted for changing industry exposure.

Finally, I chose to test the fit of the equilibrium model using the two different sets of factor loadings with another cross-sectional regression in a manner similar to Fama and MacBeth's tests of the CAPM. As described above, this test uses a cross-sectional regression on a monthly basis to measure the strength of the fit of the model. In the original study, the goals were to test certain other claims of the CAPM and to see if any other risk factors were priced besides beta. In this study, I am interested only in the fit of the equilibrium model to the returns data, and so I used a modified cross sectional model to record the R^2 statistics:

$R_i = a + \lambda_{RM-RF} b_{i_{RM-RF}} + \lambda_{SMB} b_{i_{SMB}} + \lambda_{HML} b_{i_{HML}} + \varepsilon$, where for a given month over the period 1995 – 2005:

- R_i is the return on company i
- a is an output of the regression: the average return on a stock with zero sensitivity to any of the factors
- λ_k is an output of the regression: the market price of factor k
- b_{ik} is the factor loading of factor k on company i

III. Data Used

I used the following data sources:

Monthly returns data from CRSP:[11]

- PERMNO – Unique number for each security listed by CRSP

---

[11] CRSP and Compustat data accessed through the WRDS system.

- DATE
- TICKER
- COMNAM – Company name
- EXCHCD – Exchange on which the security trades
- SHROUT – Number of shares outstanding
- PRC – Stock price
- RET – Monthly percentage return, dividend and split adjusted

Yearly segment data from Compustat:

- GVKEY – Unique number for each company listed by Compustat
- NPERMNO – Identifier that corresponds with the CRSP PERMNO for this company
- DNUM – Primary SIC code for the company
- CONAME – Company name
- SMBL – Ticker symbol
- SRCYR – Year the data was reported (can be different from the fiscal year of the company in cases like restated filings)
- SCRFYR – Month the data was reported
- STYPE – Segment type (e.g. geographical or business – I used only business segments)
- YEAR – Fiscal year of the data being reported
- FYR – Month of the end of the company's fiscal year for the data being reported
- CYR – Calendar year of the data being reported
- SALE – Segment sales for the year

- SNAME – Name of the segment
- SNAICS1 – Primary NAICS code of the segment
- SRCCYR – Calendar year of when the data was reported

Fama-French factors

- DATE – Month and year of the data
- RMRF – Return on market minus the risk-free rate
- SMB – Return on portfolio of small market cap stocks minus return on portfolio of large market cap stocks
- HML – Return on portfolio of value stocks (high book value of equity / market value of equity) minus return on portfolio of growth stocks (low book value of equity / market value of equity)

There were several issues I had with the data as it was obtained directly from the service providers. First, Compustat data for a given fiscal year is repeated if a company restates earnings. I fixed this by running a filter through the data to select only the latest source date for a given reporting period—thus I chose only the latest restated earnings. Second, Compustat segment data is reported only on a yearly basis. To match monthly CRSP returns, I split the yearly segment breakdown evenly across each month of the fiscal year reported. However, companies can change fiscal years, leaving overlapping and missing data.

To fix this, I created a new, empty table of Compustat segment data, but using a monthly basis instead of a yearly basis. Then, for each company in the data set, for each reported year of the company's segment data, starting at the earliest and ending with the latest year's data, I computed the total sales for that company for that fiscal year. I did not include corporate segments that have negative sales. I computed each segment's percentage of the total sales based on this modified total sales number. Then, starting with the last month of that fiscal year and working back to the first month of the fiscal year, I copied the segment percentage data into the new data table for each month. Last, I iterated through all months of the new data table reported for the company in order, and if a month was missing, I copied the previous month's data forward.

This procedure had several effects. First, in the case of overlapping fiscal reporting periods, it ensures the newer data takes priority. Second, in the case of missing months, the latest data from previous months is copied forward to fill the gap.

IV. Procedure

I selected a set of companies to examine for which I had enough data: at least eleven years of both returns and segment data (or 132 monthly observations): that is, beginning on or before January 1995 and ending on December 2005. This resulted in a set of 1,994 companies.

I then computed a set of factor loadings for each company over the time period studied, using the traditional regression technique. The model I used was the factor model described above as specified by Professors Fama and French.

To do this, I exported each series of company returns into the statistical package R and ran the APT regression in order to determine each of the factor loadings (*b*'s). I exported the resulting factor loadings for each company back into my database. Since these factor loadings are estimated over the entire period I studied and do not change over the period, I will refer to them as the "static" factor loadings.

I examined the industries reported by the companies in the set I had selected. I used the first two digits of the SNAICS1 field from the Compustat data to group industries. There were 26 of these industry groups, including a “00” group whose Compustat data were empty and a “99” group, which Compustat uses to indicate a non-operating or liquidating company.

Then I created a set data matrices that could be used by my statistical package for performing the regression to derive returns series for each industry. I created one such matrix for each month in the period I studied. Each matrix contained a cross-sectional set of data for each company listed in Compustat and CRSP (in order to maximize the number of observations for calculating industry returns). The columns of the matrix contained the company identifier, the company’s return, and columns containing the company’s exposure to each industry, as shown in the example column headers below:

PERMNO	RET	Industry 1 pct.	Industry 2 pct.	…	Industry *N-1* pct.

I discarded industry percentage columns for which the sum of the squared values equaled zero (to eliminate industries that had no corresponding companies in that month), and dropped the column with the lowest sum of squared values. The second modification ensured that the data fit the statistical model described above. I then exported the corresponding data matrix into the statistical package R, and ran a cross-sectional regression to derive each industry’s average return in that month.

The regression resulted in coefficients that represented the difference in returns between each industry and the industry whose column I dropped from the data matrix. To correct for this, I took the constant term to be the return on the industry I had dropped, and added that constant term to each coefficient term in order to derive each industry’s returns for that month. I exported these returns back into my database. I repeated this process for each month from January 1995 to December 2005.

Once I had derived returns for each industry for the entire time period, I computed a set of factor loadings for each industry using the derived industry return data. I used the same Fama-French APT model and procedure as I had used in computing each company's factor loadings, except instead of company returns, I used the derived industry returns. I exported the resulting factor loadings for each industry back into my database.

Based on these industry factor loadings, I calculated an alternate set of factor loadings for each company for each month. As described above, I calculated the factor loadings on a portfolio of securities as weighted averages of the factor loadings on the individual securities in the portfolio, weighted by the portfolio weight of each security. Therefore, in each month, I calculated the factor loadings of a security by weighting the factor loadings of each industry by the company's exposure to each industry (percentage of total sales), and adding them together.

This procedure ensured that *industry* factor loadings do not change over time in this model. However, *company* factor loadings change when their reported industry exposure changes. I will refer to these factor loadings as the "dynamic" factor loadings, as opposed to the traditionally calculated "static" factor loadings calculated previously.

Then I ran a series of cross-sectional regressions to test the strength of this method of calculating factor loadings. For each month in the period I studied, I created two tables, each containing cross-sectional data for every company in the sample. The first table contained each company's return and factor loadings as calculated by a traditional regression technique. The second table contained each company's return and factor loadings as calculated by the derived-industry-returns technique. I exported these tables to R and performed the cross-sectional regression described earlier. I then saved all the regression statistics from R back into my database for later comparison.

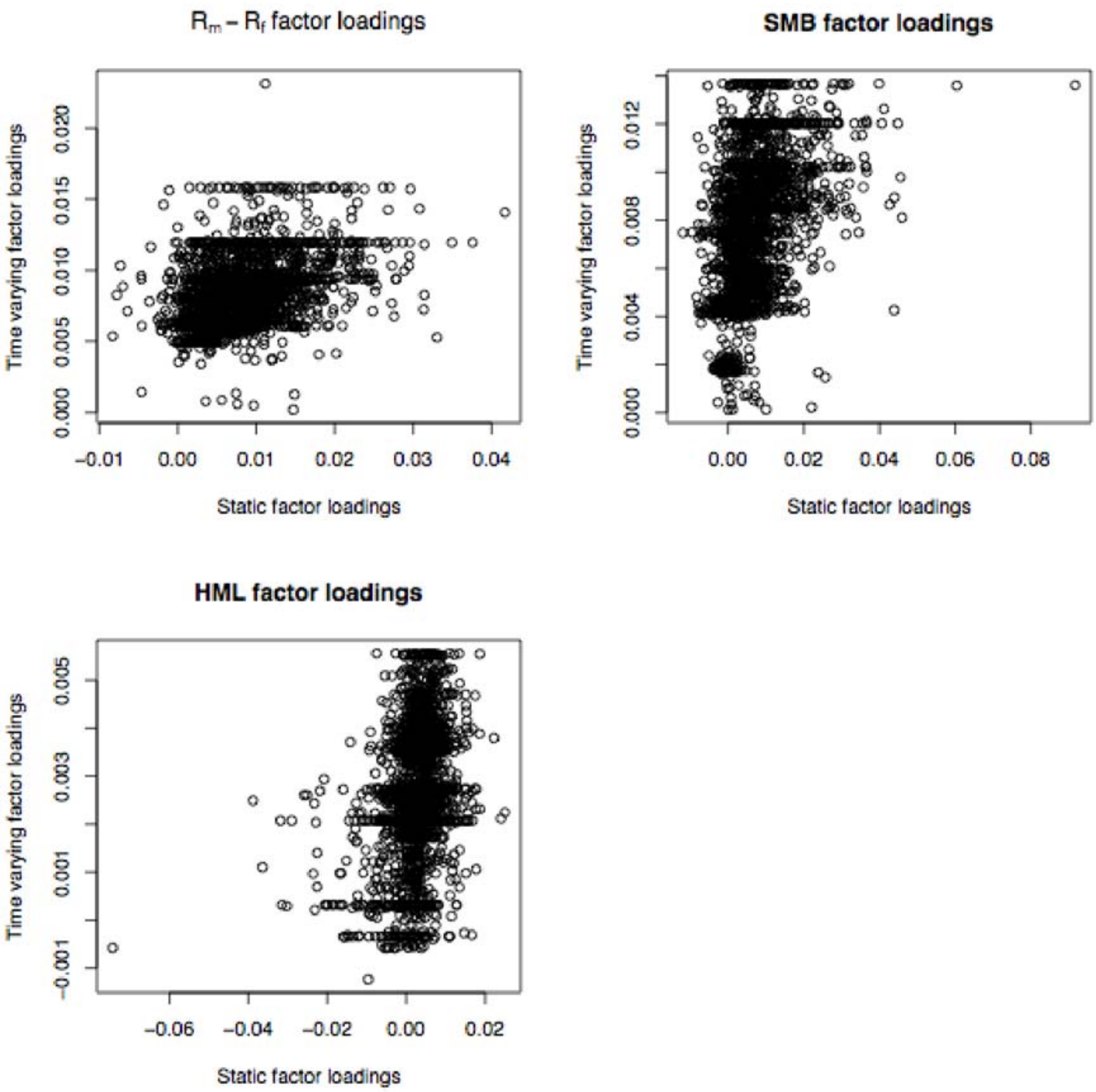
Because of the results that are described below, I did further cross-sectional tests. These followed the same procedures as in the previous paragraph, but instead of including all 1,994 stocks, I took sub-samples of the companies for each other test. I examined the subset of those 1,994 stocks that had reported sales from more than one distinct industry over the entire time period, and another subset of stocks that had reported sales from more than two distinct industries to test whether the number of industries a company is exposed to influences these regressions.

I also attempted to derive a measure of how much a company changes over time. For a given company and industry, I took the difference between the percentage of sales in a particular time period and the percentage of sales in the previous time period as a measure of how that company had changed in a particular industry in a particular month. I added the absolute value of this measures across all industries and across all time periods studied for each company. I took this as the overall measure of how much a company's industry exposure changed: it has a bottom limit of zero if the company did not change at all and no practical upper limit. I divided the group of 1,994 companies into two equal-sized groups: those that had a low change measure and those with a high change measure. I ran the cross-sectional regression on these two last, expecting that the high-change group would show better results than the low-change group.

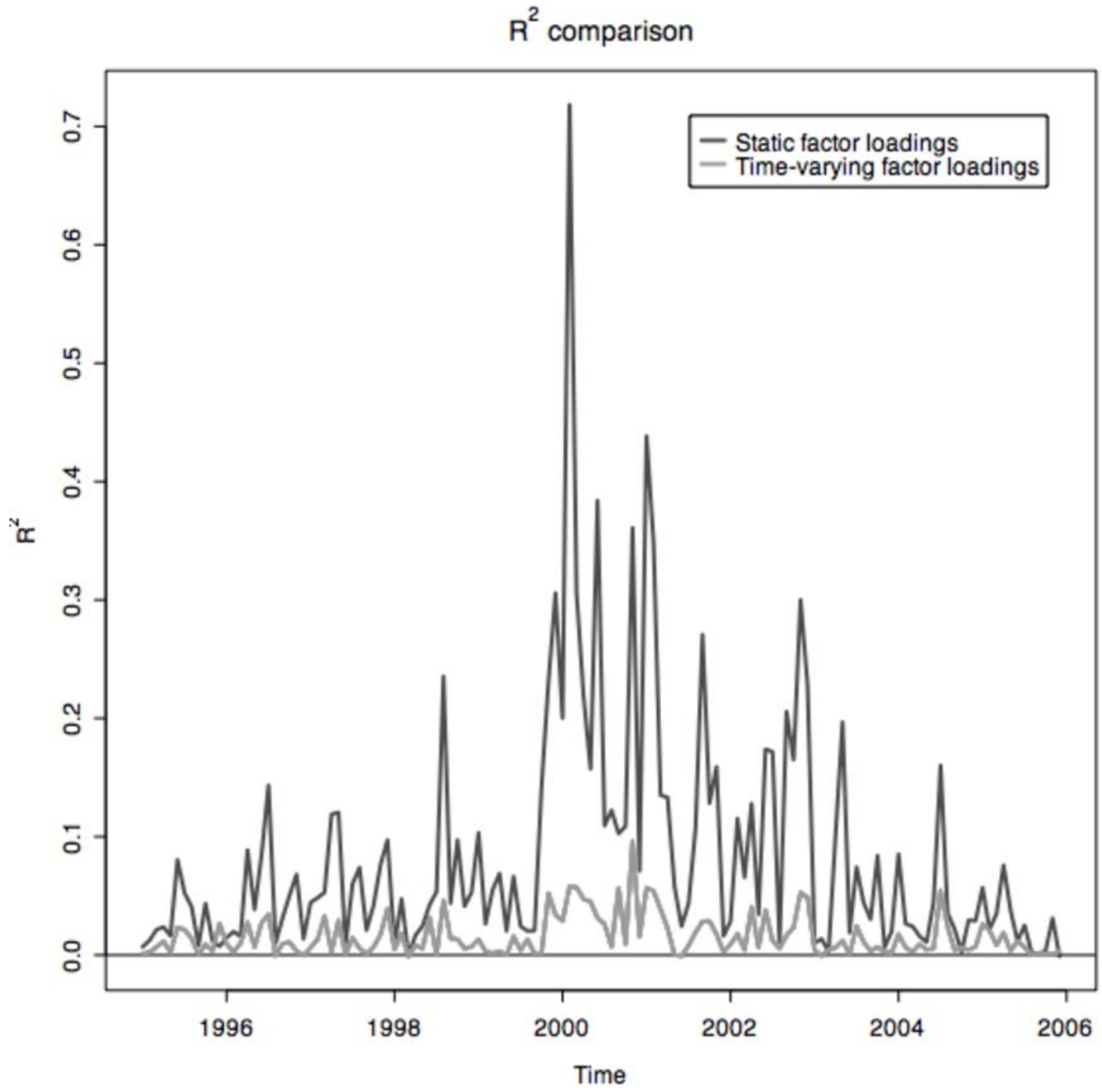
V. Results

Overall, the data show a complete lack of support for the original hypothesis. The R^2 measures of the regressions show that these factors explain very little of the variation in the stock returns. More importantly, the tests using the dynamic factor loadings that were the main point of the study had a significantly lower average R^2 than the tests using the traditionally calculated, static factor loadings. Even the tests of subsets of the companies showed similar results.

The following charts show scatterplots for each of the three sets of factor loadings, of the static version versus the dynamic (or time-varying) version. The dynamic factor loadings in these charts are averages of each factor loading across all time periods for each company. When creating these charts, I did not expect to see any particular relationship between the two types of factor loadings: after all, I am trying to improve on the estimation of this measure. However, there are several interesting points to observe. First, there are several horizontal bands in each chart. I believe this represents companies who report the same industry, but have wildly different factor loadings. To extend Professor Damodaran's "bottom-up beta" analogy, there may be factors other than a company's industry exposure that influence that company's exposure to a particular risk factor; this analysis does not capture them. Second, notice the scales of the three charts. The static factor loadings for the excess market return $R_m – R_f$ range from around -0.01 to a little over 0.04; the corresponding dynamic factor loadings range from 0 to just above 0.02. The disparity between the scales of the two types of factor loadings is even greater for the other factor loadings. This seems to show that many of the companies in the sample have greater exposure to these risk factors than indicated by the weighted average of their component industries' exposure to these factors.



The following chart shows the actual results of the cross-sectional regression meant to test the strength of this industry-beta methodology. It depicts two time series of adjusted R^2s, output from each of the monthly cross-sectional regressions I performed. Note that the actual R^2s of these regressions were not materially different from the adjusted R^2s. The solid black line represents the regressions using the traditionally calculated static factor loadings, while the gray line represents the regressions using the dynamic factor loadings. As shown in the chart, except for a few months, the R^2s of the regressions using the static factor loadings was significantly greater than those of the regressions using the dynamic factor loadings. There is no other conclusion to be drawn except that the dynamic factor loadings as I calculated them are poorer estimates of the "true" factor loadings than the static factor loadings, directly contradicting my hypothesis.



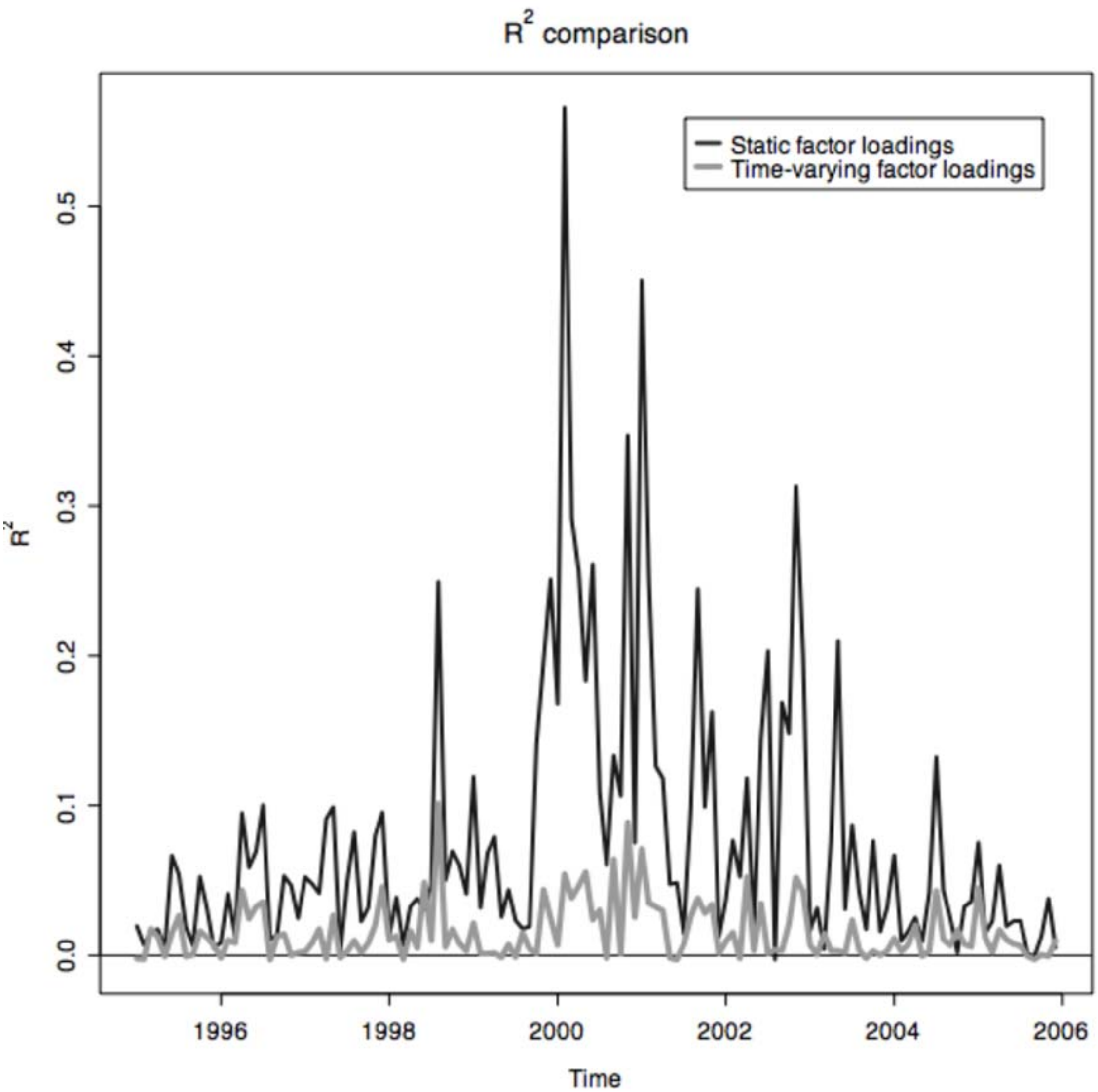
In addition to this graph, I found an average of the R^2s for the regressions using each set of factor loadings. For the 132 monthly regressions using static factor loadings, I found that the average R^2 (*not* adjusted R^2) from January 1995 to December 2005 was 8.6%. For the corresponding regressions using dynamic factor loadings, the average R^2 was 1.7%. For reference, *Modern Portfolio Theory and Investment Analysis* by Elton, Gruber, Brown, and Goetzmann list R^2s for Fama and MacBeth's cross-sectional tests of the CAPM equilibrium model. Their study used 20 beta-ranked portfolios of securities, rather than individual securities, to minimize the beta estimation error, so the R^2s they report are not directly comparable to those I found. However, as a reference, they report an R^2 of 29% for their basic CAPM test over the period 1935 – 1968.[12]

Faced with these disappointing results, I attempted to see whether there was any glimmer of hope for my hypothesis. Since this methodology was used successfully in the Japanese markets, perhaps it works better for companies that are in more than one line of business—conglomerates. Alternatively, as explained above, the Japanese phenomenon may have been due more to the amount of change in industry structure rather than the number of cross holdings in each company. To examine this possibility, I ran the cross-sectional regressions again on subsets of the 1,994 companies.

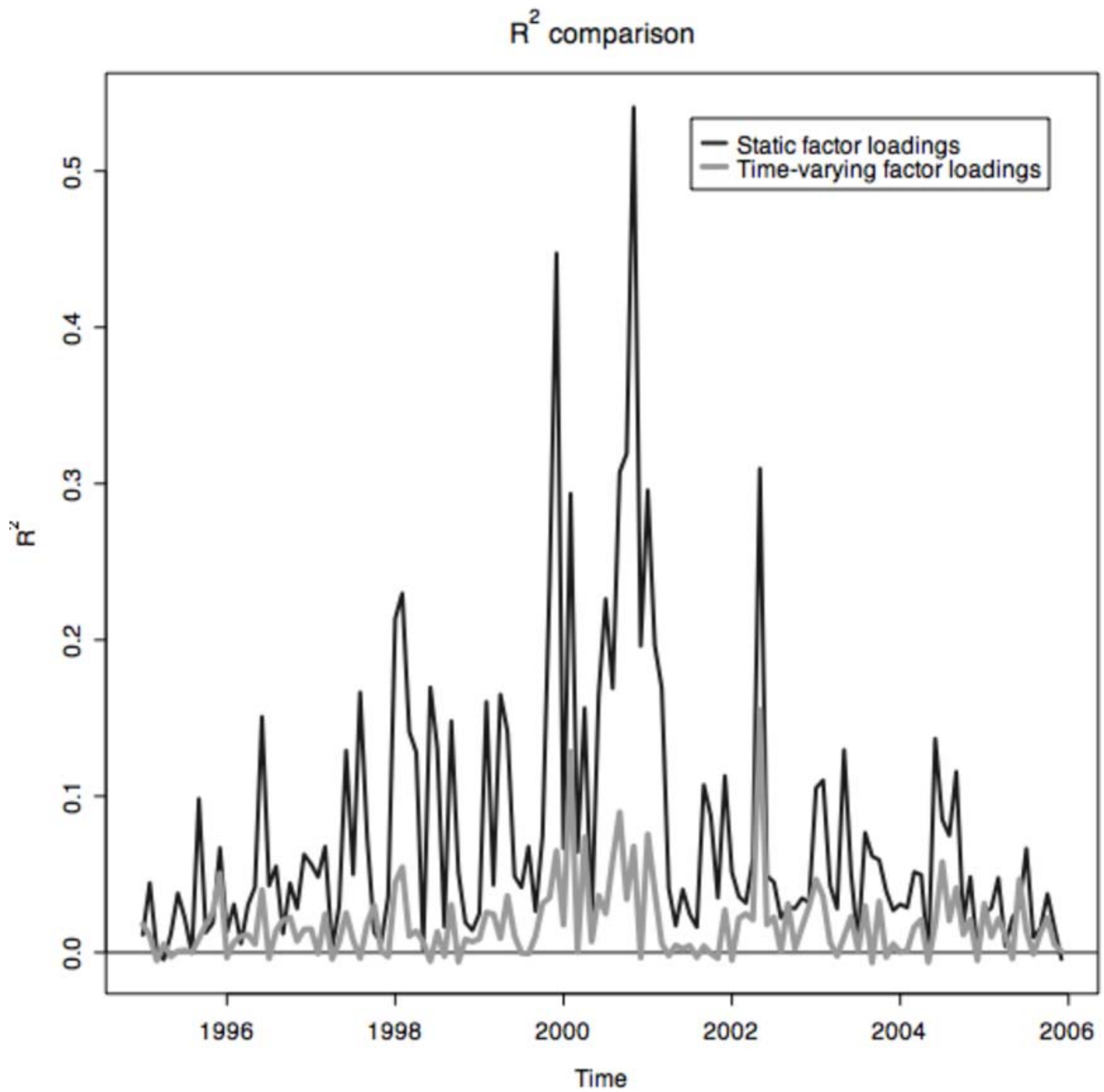
First, I examined subsets based on how many industries each company was in. I selected subsets of the companies that, at any point over the 132 months I examined, had reported sales from more than just one segment, narrowing the number of companies to 987. Next, I selected subsets of the companies that had reported sales from more than *two* segments, further narrowing the number of companies to 447. The corresponding adjusted-R^2 plots of the resulting cross-sectional regressions are given below. First, the companies with more than one segment:

---

[12] Elton et al. "Chapter 15: Empirical Tests of Equilibrium Models." *Modern Portfolio Theory and Investment Analysis, Sixth Edition*. John Wiley & Sons: 2003. p. 348.



Next, the companies with more than two segments:



As shown, these results are hardly different; they lead to the same conclusions as the regression over all 1,994 stocks. The average R^2s were:

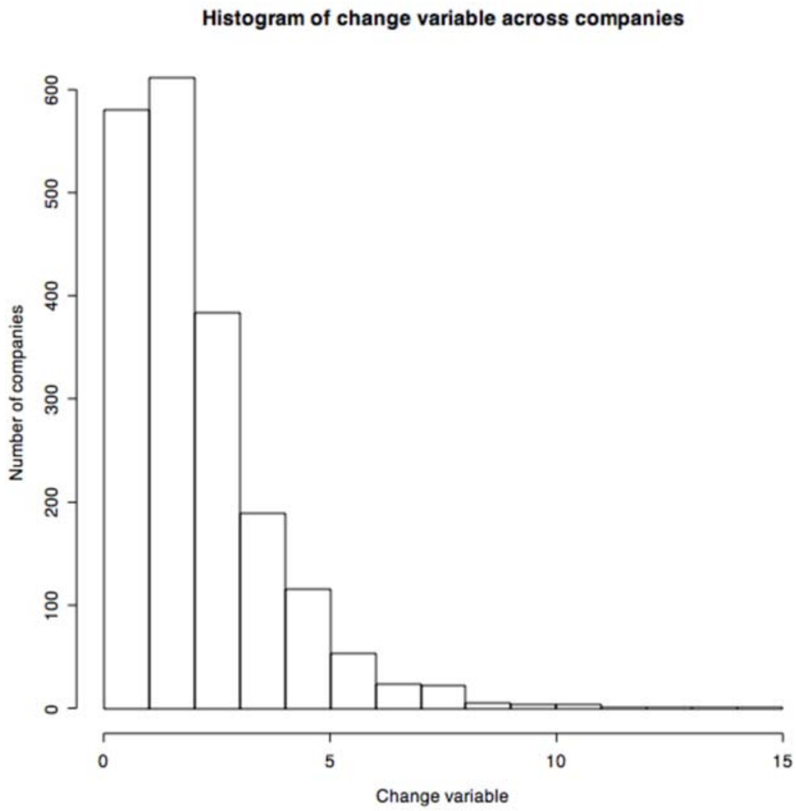
Companies with:	Dynamic factor loadings	Static factor loadings
> 1 segment	1.9%	8.1%
> 2 segments	2.5%	8.4%

These results show that the dynamic factor loadings are no better a representation of the factor loadings for companies with more than two segments than they are for companies with more than one segment.

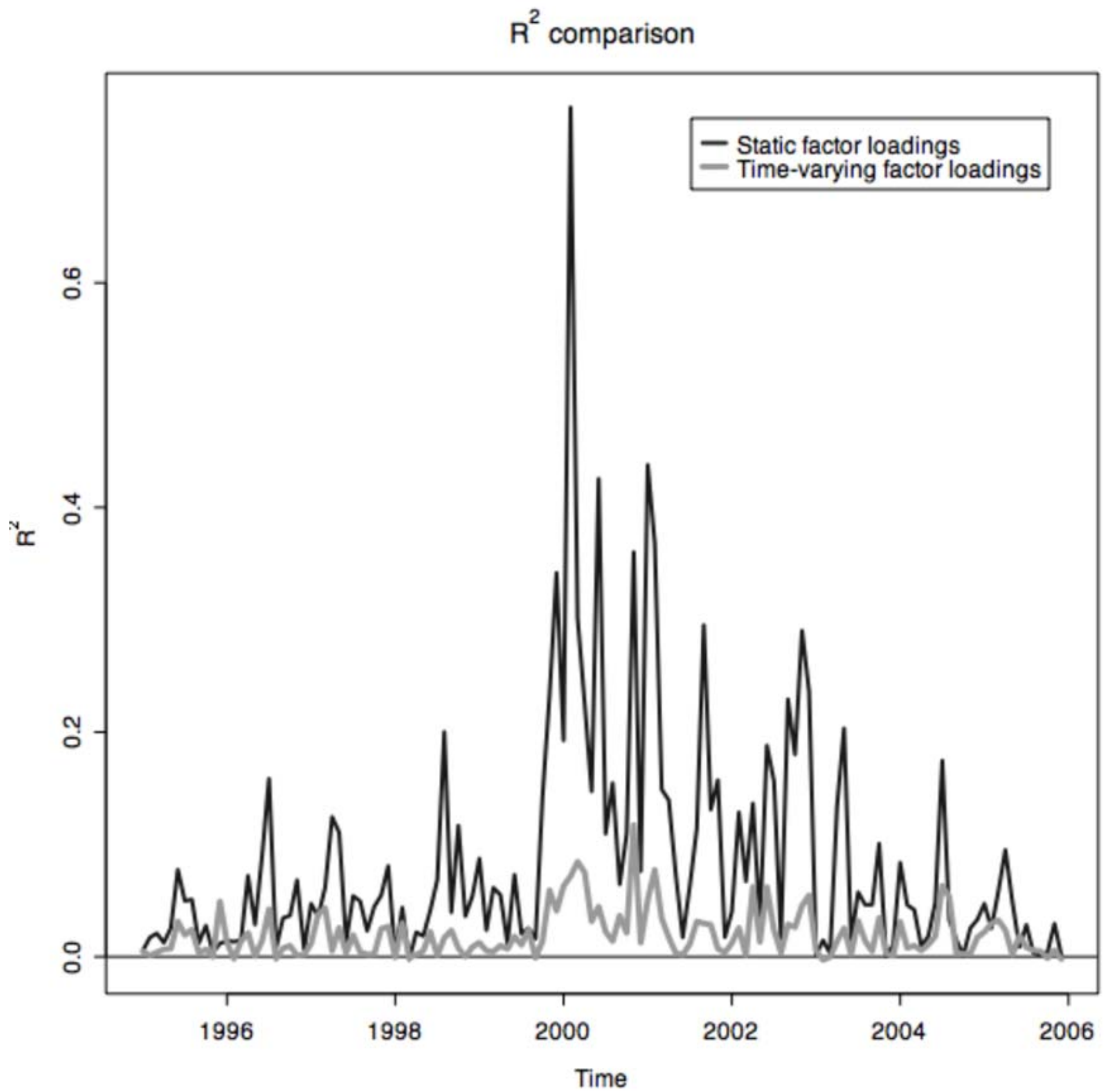
What about the degree of change of exposure within a company? First, I needed to estimate this figure. I decided to estimate it by computing the change in percentage-of-sales for each company, for each industry, from one time period to the next. Since the sum of a company's percentages of sales across each industry must sum to 100% for any given time period, the sum of the change in the company's percentage of sales across all industries must also sum to 100% for a given time period: a company can shift its industry exposures from one industry to another, but the addition to one will be exactly offset by a decline in the other in percentage terms. Therefore, for each company, I added the *absolute value* of the computed change across all industries and across all the time periods I studied. This gave a numerical figure representing the degree to which a company changed its industry exposures, with a lower bound of zero for a company that did not change its industry exposure at all, and an upper bound limited by the number of periods I studied. The highest change score among the 1,994 stocks was 14.26, for US Energy Corp (ticker USEG), a company that has been involved in Mining, Minerals, Commercial Operations, Retail Sales, Oil & Gas, and Construction Operations over the course of its history as reported by Compustat.

The histogram of this change variable is given below:

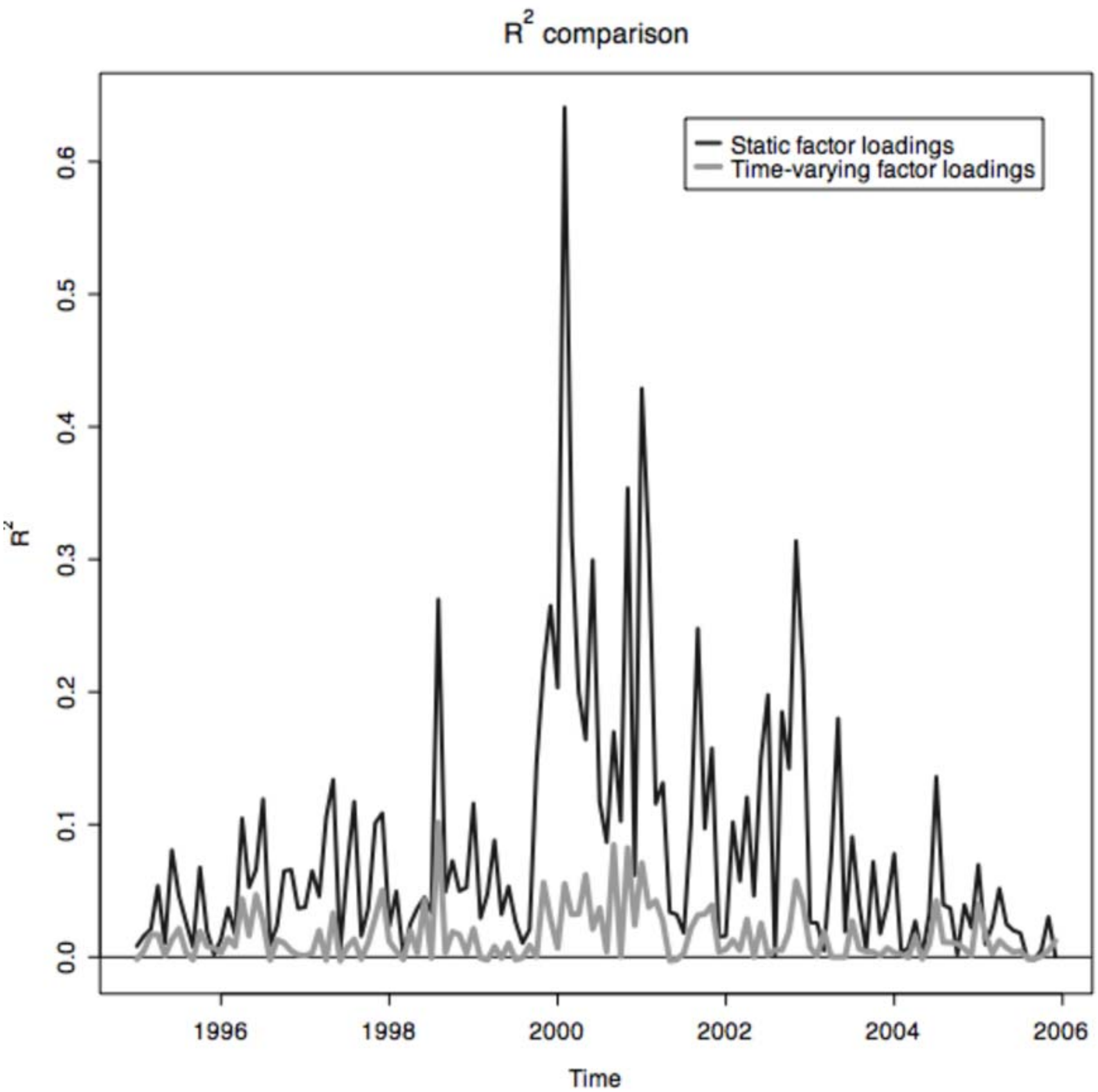
Measure	Value
High	14.26
Low	0.00
Mean	2.07
Median	2.00
Std Deviation	1.79
Skewness	1.40
Kurtosis	7.38



I split the set of companies into two equal-sized groups ranked by the change measure and performed the cross-sectional tests on these two groups. The low-change group had an average change measure of 0.866; the high-change group had an average change measure of 3.279. The corresponding plots of the adjusted R^2s are given below. First, the low-change group:



Next, the high-change group:



These plots again show that the dynamic factor loadings are significantly worse than the static factor loadings. Even more troublesome, the R^2s do not improve for the high-change group, either alone or in comparison to the low-change group. An average of the R^2s for each test is shown below:

Companies with:	Dynamic factor loadings	Static factor loadings
Low change	2.2%	8.9%
High change	1.9%	8.5%

Despite my hypothesis that the amount of change in industry composition will determine whether these dynamic factor loadings will make a difference, I found no evidence even of an improvement in fit for companies that exhibited a higher degree of change over those with a lower degree of change.

VI. Conclusion

Unfortunately, I found no evidence to support my hypothesis. While I predicted that the R^2s of the cross-sectional regressions using the dynamic factor loadings would be significantly greater than those of the regressions using the static factor loadings, they were in fact significantly lower.

There are several possible ways to explain the results I obtained. The first and foremost possibility is that I made one or more errors in my analysis. This could range from something as fundamental as a conceptual idea that I missed to procedural errors, mistyped commands, and data problems I did not address. Alternatively, my hypothesis could just be plain wrong.

A second possibility is that the data are of insufficient quality to support my analysis. The Compustat segment data are reported on a yearly basis, while I am studying monthly returns, a mismatch that forces an estimation procedure in order to proceed with a monthly analysis. Also, the Compustat segment data are self-reported, causing many gaps in the data; potentially, changes in reported segmentation could be reported while the underlying industry exposure remains the same.

Third, the time period I studied, 1995 to 2005, did not see much empire building or destroying activities. It may be interesting to repeat this analysis for periods of higher industry change, like the formation of conglomerates in the 1960s or the break-up of such companies in the 1980s.

Fourth, there was some indication of poor regression fit during the computation of industry returns. As mentioned above, I performed a series of cross-sectional regression of company returns against industry exposure for each month studied. While checking the regression statistics I noticed that the variance inflation factors (VIF) of the first month's regression coefficients were extremely high, even after using the modified model, indicating that there was a high degree of multicollinearity between the independent variables. This means that the regression is unstable: if any of the independent *b* variables changed only slightly, the fitted regression coefficients would change dramatically, indicating that the industry returns derived from these regressions could have been estimated with significant error.

Lastly, my analysis makes the assumption that company returns, on average, are composed of industry returns, weighted by a percentage-of-sales measure. However, this may be an inappropriate methodology for calculating returns, and may have caused errors throughout the analysis.

There are several steps that could be performed to enhance this analysis and potentially obtain evidence in support of the hypothesis. First, the data should be cleaned up, with gaps in reported segmentation closed, dramatic changes in segmentation checked against regulatory filings, and any other data issues resolved. Of course, this would be a very difficult task given the amount of data I used, so a subset of the data may need to be used. Second, as mentioned earlier, tests of equilibrium models traditionally use portfolios of securities to minimize estimation error. Repeating the cross-sectional tests of this study using portfolios of securities may be a better formal test of the hypothesis.