TOWER BUILDING AND STOCK MARKET RETURNS

Gunter Löffler*

University of Ulm

April 2013

Abstract

Construction starts of skyscrapers predict subsequent US stock returns. The predictive ability exceeds that of alternatives such as the prevailing historical mean, predictions based on dividend ratios and recently suggested combination forecasts. One explanation for these patterns is that tower building is indicative of over-optimism; alternatively, tower building could help to identify periods of low risk aversion. I present indirect evidence that is consistent with both explanations.

JEL classification: G12, G14

*I am very grateful for comments received from the referee and the editors as well as from Erik Theissen and conference and seminar participants in Hamburg (German Finance Association), Hannover (Leibniz University) and Rome (IFABS).

Gunter Löffler, Institute of Finance, University of Ulm, Helmholtzstrasse 18, 89069 Ulm, Germany. Phone: ++49-731-5023597. Fax: ++49-731-5023950. E-mail: gunter.loeffler@uni-ulm.de

I. Introduction

Ever since the story of the tower of Babel was recorded, the construction of large towers has been associated with human hubris. From a finance theory perspective, towers are large-scale projects with uncertain future cash flows and large funding requirements. These observations suggest two reasons why tower building might predict low future stock market returns. Either it indicates periods in which over-optimism has led to overvalued stock markets, or it helps to identify times of low risk aversion. (With low risk premia, funding costs for large-scale projects are lower, while future stock market returns are expected to be relatively low as well.) An example that illustrates both interpretations is the Chicago Spire, which had a planned height of 609 meters.[1] Construction of the Chicago Spire began in June 2007, a time in which (i) risk premia – as exemplified by low credit spreads – were low, and (ii) valuation levels appear to have been relatively high. Though the Chicago Spire stands out because of its planned height, it is representative of the many high-rise buildings planned that year. The number of towers taller than 100 meters that were started to be built in 2007 was more than twice the annual average of such construction starts over the 20 years from 1987 to 2006.[2]

In this paper, I therefore examine whether tower building is associated with lower subsequent stock market returns. Building activity is measured through construction starts of towers that exceed a trailing mean tower height. In the US, the predictive power of this measure compares favorably to the predictive power of the dividend price ratio, a variable that has been studied extensively in the literature (e.g. Welch and Goyal, 2008), as well as to recently suggested combination forecasts (Rapach, Strauss and Zhou (2010) and Ferreira and Santa-Clara (2011)).

[1] Construction of the building was halted in 2008.
[2] Data are from Emporis, which is described in Section II.

Further analysis shows that international tower building activity predicts a world ex US stock market index.

The two possible explanations for the predictive content of tower building are difficult to separate empirically. Indirect evidence is consistent with both explanations. With rational asset pricing, forcing return forecasts to be non-negative should increase predictive accuracy, but the effects of such a constraint are mixed. Furthermore, both credit market conditions and sentiment variables explain construction starts of large towers.

The perception that tower building can be linked to economic as well as stock market performance is frequently voiced in the media.[3] Often, news articles cite the research report of Lawrence (1999) and follow-up reports, e.g. Lawrence, Hsu, Luo, and Chan (2012). The only associated academic paper I found is Thornton (2005), who discusses the relationship between tower building, business cycles and economic crises but does not conduct a statistical analysis. Barr (2010) empirically examines the determinants of skyscraper height and concludes that status plays a role, leading to heights that exceed the profit maximizing height.

There is a large body of literature on predicting stock markets with dividend ratios and other variables. Classical references are Campbell and Shiller (1988) and Fama and French (1988). Recent contributions include Goyal and Welch (2003), Malkiel (2004), Fisher and Statman (2006), Boudoukh, Richardson and Whitelaw (2008), Campbell and Thompson (2008), Cochrane (2008), Welch and Goyal (2008), Rapach, Strauss and Zhou (2010) and Ferreira and Santa-Clara (2011).

---

[3] Examples include a 2005 article in *Fortune* (http://money.cnn.com/magazines/fortune/fortune_archive/2005/09/05/8271392/index.htm) and a 2009 article in *The Telegraph* http://www.telegraph.co.uk/news/worldnews/middleeast/dubai/6934603/Burj-Dubai-The-new-pinnacle-of-vanity.html)

Whether stock market returns can be predicted is still controversial. While the evidence for in-sample predictability appears strong, out-of-sample evidence is much weaker. A possible reason for this wedge is structural breaks in fundamentals (Lettau, Ludvigson and Wachter (2008) and Freeman (2011)).

II. Data and Methodology

Data on towers are obtained from the research database of Emporis, a private information provider focusing on building-related information.[4] The database also contains information on planned projects and construction status. I include buildings that were started to be built but were never finished, thereby avoiding a possible selection bias that might arise if only finished buildings were studied. The measure of height used is the elevation from the base of building to its highest architectural element.[5] I select buildings that are classified as either skyscrapers or high-rise buildings in the Emporis database, and further exclude city halls, county halls, capitols and courthouses. The reason for excluding public buildings is that linkages to the stock market are likely to differ. Government building activity should be less sensitive to market conditions, and may even be countercyclical if governments invest in buildings to smooth business cycles. The procedure also excludes airport towers, chimneys, churches, telecommunications towers and other structures that are not classified by Emporis as skyscrapers or high-rise buildings. For simplicity, I will henceforth use the word *tower* for private high-rise buildings that have been selected according to the procedure described above.

---

[4] The quality of the database was confirmed through cross-checks with Condit (1964), Landau and Condit (1996) as well as official websites of existing buildings.

[5] "Architectural elements include everything which is integral to the design, including sculptures, spires, screens, parapets, and decorative features." (Source: http://standards.emporis.com).

In the definition of variables, I focus on construction starts rather than on construction completions because construction starts should provide a better measure of the current situation – be it overvaluation or low risk premia. Several towers like the Chrysler Building and the Sears Tower were completed despite the fact that economic conditions worsened significantly after their start. In unreported analysis I re-defined the variables using the completion dates and found that the predictability is lower.

As a measure of US tower-building activity, I examine the number of towers exceeding a threshold defined by the trailing average height of large buildings. Since the height of large towers is trending upward over time, this appears superior to the use of a fixed threshold such as 100 meters. Specifically, I define the threshold to be the average height of buildings of over 50 meters that were built in the thirty-year interval before the year in question. The choice of thirty years is motivated by the observation that building activity can follow relatively long cycles. For example, in the 20 years from 1933 to 1952, only 76 towers of over 50 meters were built, and these had an average height of 80.1 meters. This compares to an average height of 88.9 meters for the 473 towers whose construction began during the 1923-1932 period. Using a trailing average of thirty years makes the threshold less dependent on general cycles in building activity, and thus better suited to identifying buildings that would be considered tall.

Together with general building activity, the number of towers rising above a given threshold is also trending upward, even if the threshold is time-varying. I therefore examine logarithms of US building counts relative to their 30-year trailing average. The precise definition of this variable *LargeStart* is given below:

Auxiliary variable:

$$\text{THRESHOLD\_b}_t = \text{Average height of towers taller than 50 meters, } \textit{begun} \text{ in t-1, t-30}$$

Measure of tower-building activity:

$$\text{Largestart}_t = \ln\left(\frac{1 + \text{Number of towers with height above THRESHOLD\_b}_t \textit{ begun} \text{ in year t}}{1 + \text{Average annual number of tower starts above THRESHOLD\_b over t-1, t-30}}\right)$$

Figure I shows the building activity as measured by *LargeStart*.

[INSERT FIGURE I HERE]

Stock market data are obtained from different sources. Annual US stock market returns, associated dividend information, and risk-free returns for the time period 1871 to 2010 are obtained from Robert Shiller's website.[6] For the 1926-2010 sub-period, annual data on the value-weighted market portfolio and dividend information are from CRSP, made available by Michael Roberts;[7] the risk-free rate that is used for the CRSP data is the one-month treasury bill rate taken from Ken French's website.[8]

In the literature on stock market predictability, dividend-based ratios are the most widely studied predictors for long-horizon stock market returns. I follow the literature and use the dividend price ratio, defined as the logarithm of dividends paid over the last year divided by the current index value; it will be denoted *dp*. Other predictors, including combination forecasts, as

---

[6] http://www.econ.yale.edu/~shiller/data.htm

[7] http://finance.wharton.upenn.edu/~mrrobert/data_code.htm

[8] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/

suggested by Rapach, Strauss, and Zhou (2010) and Ferreira and Santa-Clara (2011), will be studied in the sensitivity analysis of Section IV.

Stock market returns enter the analysis as logarithmic excess returns over the risk-free rate, denoted by $r_{t,t+k}$. In using log returns, I follow Fama and French (1988) and Welch and Goyal (2008). One justification is the fact that log returns are closer to being normally distributed, which should increase the reliability of regression analysis.[9]

In line with the extant literature, predictability is analyzed through linear regressions. Information contained in tower-building activities may be reflected with a time lag, but the variable capturing the building of large towers is highly autocorrelated by construction,[10] which is why I do not include further lags. Thus regressions are of the form:

$$r_{t,t+k} = b_0 + b_1 dp_t + b_2 LargeStart_t + u_{t,t+k} \qquad (1)$$

For return horizons larger than one year (k>1), the return observations are overlapping, which induces correlations in the error terms. While ordinary least squares regression continues to yield consistent estimates of the coefficients *b*, standard errors are no longer reliable. Until recently, the common academic response was to use Newey and West (1987) standard errors. Ang and Bekaert (2007), however, have shown that the Newey and West procedure is sensitive to time persistence in the explanatory variables. Through simulations, Ang and Bekaert (2007) show that Hodrick (1992) standard errors are much more reliable in regressions such as (1). With k>1, I

---

[9] For the 1871-2010 data, regressing 5-year log returns on the dividend price ratio leads to residuals whose normality is not rejected by a skewness/kurtosis test (p-value = 0.615); with simple returns, normality of residuals is rejected (p-value=0.011).

[10] The first-order autocorrelation is 0.82. In unreported analysis, I added a one-year lag of LargeStart. Across the specifications of Table 1, its coefficient was not significant.

therefore use Hodrick (1992) standard errors. With one-year returns, the standard heteroscedasticity-robust estimator of White (1980) is used.

I also will explore the out-of-sample performance of predictions based on regressions such as (1). One of the metrics examined is the out-of-sample R² suggested by Campbell and Thompson (2008), which compares the mean squared error of a prediction model to the errors one would incur when using the historical mean prevailing at time t as a predictor. Let $\bar{r}_{t,t+k}$ denote the prediction derived from the historical mean prevailing at time t, and let $\hat{r}_{t,t+k}$ denote an out-of-sample regression-based prediction. The out-of-sample R² is then computed as follows:

$$R_{OS}^2 = 1 - \frac{\sum_{t=m}^{T} \left(r_{t,t+k} - \hat{r}_{t,t+k} \right)^2}{\sum_{t=m}^{T} \left(r_{t,t+k} - \bar{r}_{t,t+k} \right)^2} \tag{2}$$

where m is the starting year of the out-of-sample analysis. If the R_{OS}^2 is positive, the prediction model outperforms the prevailing mean, which serves as a natural benchmark for evaluating predictive performance. To assess the statistical significance of the out-of-sample R², I follow Rapach, Strauss, and Zhou (2010) and favor the Clark and West (2007) MSPE-adjusted statistic over the Diebold-Mariano (1995) test, which can have low power when applied to nested models.[11] To compute the MSPE-adjusted statistic, define

$$f_{t,t+k} = \left(r_{t,t+k} - \bar{r}_t \right)^2 - \left(\left(r_{t,t+k} - \hat{r}_{t,t+k} \right)^2 - \left(\bar{r}_{t,t+k} - \hat{r}_{t,t+k} \right)^2 \right), \tag{3}$$

[11] See the discussion in Rapach, Strauss, and Zhou (2010). The prevailing mean is nested in regression models of type (1) because it would result from restricting coefficients other than b_0 to be zero.

and regress $f_{t,t+k}$ on a constant. The p-value for a one-sided test is obtained by applying the standard normal distribution to the t-statistic of the constant. For k>1, I use Newey and West standard errors with lag k.[12]

III. Tower Building and Stock Market Returns

In-Sample Analysis

To begin, I regress future stock market returns on the variable *LargeStart*. This specification, which I will call the tower model, is compared to a regression with just the dividend-price ratio (the dividend model), as well as to a regression that includes both variables (the dividend+tower model). Specifications are thus:

Dividend model: $$r_{t,t+k} = b_0 + b_1 dp_t + u_{t,t+k} \tag{4}$$

Tower model: $$r_{t,t+k} = b_0 + b_1 \text{LargeStart}_t + u_{t,t+k} \tag{5}$$

Dividend+tower model: $$r_{t,t+k} = b_0 + b_1 dp_t + b_2 \text{LargeStart}_t + u_{t,t+k} \tag{6}$$

Table 1 presents the results for prediction horizons of one, three, and five years, separately for the 1871-2010 and 1926-2010 samples. As is familiar from the literature, high dividend price ratios are associated with high future returns. However, the coefficients are at best marginally

[12] Note that the doubts about the reliability of Newey and West (1987) discussed in conjunction with regression (1) arise from the persistence of the predictors, and therefore do not carry over to the regression that is run for the MSPE-adjusted statistic.

significant, which is in line with the findings of Ang and Bekaert (2007).[13] The tower model, by contrast, shows significant predictive ability across the horizons. High tower building activity goes along with low future returns. Coefficients are statistically significant on a 5% level or better, and they are economically significant as well. The variable *LargeStart* has a standard deviation of 1.05. Together with the coefficient of -0.112 in the three-year regression, for example, this implies that a one standard deviation increase in tower building activity lowers expected three-year returns by 11.8 percentage points. This translates into a per annum return difference of 3.9 percentage points. The dividend+tower model regressions confirm the results. The dividend-price ratio remains insignificant once tower building is controlled for. *LargeStart* remains significant on the 5% level except for one regression, in which the t-statistic is -1.957.

[INSERT TABLE 1 HERE]

Out-of-Sample Analysis

An out-of-sample analysis mimics the situation of market participants who tried to use the information in predictive variables at a certain point τ in the past. To predict *k*-year returns starting at the end of year τ, one would run the regression (when using the tower model)

$$r_{t,t+k} = b_0 + b_1 LargeStart_t + u_{t,t+k}, \quad t = 1,...,\tau - k, \tag{7}$$

derive coefficient estimates $\hat{b}$, and compute the prediction

$$\hat{r}_{\tau,\tau+k} = \hat{b}_0 + \hat{b}_1 LargeStart_\tau \tag{8}$$

[13] With the Newey and West estimator, the coefficient on the dividend yield has a t-statistic of 2.48 in the five-year regression.

Predictions for the other two models are derived in the same fashion. By running regressions of type (7) for each τ considered in the analysis, one obtains a series of out-of-sample predictions. Following the suggestion of Campbell and Thompson (2008), I also examine the effects of imposing a non-negativity constraint on the prediction. The motivation is that expected excess returns on an asset exposed to systematic risk should be non-negative if the average investor is risk-averse. Since the returns $r_{t,t+1}$ are log returns, negative forecasts can be consistent with rationality as the relevant simple returns also depend on volatility. Assuming normally distributed log returns and estimating future volatilities with the standard errors of the predictive regressions leads to the following predictor that constrains predicted simple returns to be non-negative:

$$\hat{r}_{\tau,\tau+k}(constrained) = \max\left(-\hat{\sigma}_{\tau}^{2}(u_{t,t+k})/2,\, \hat{r}_{\tau,\tau+k}\right), \qquad (9)$$

where $\hat{\sigma}_{\tau}(u_{t,t+k})$ denotes the estimated standard error of a regression that is run to make a prediction from time τ.

The first date τ on which a prediction is made is chosen to be 1910. By then, there are least 35 observations for the regressions on which the predictions are based.

Out-of-sample performance will suffer if coefficient estimates are unstable. It is, therefore, illustrative to examine how coefficients change over time. Figure II shows the time series of slope coefficients for the dividend model and the tower model. Slope coefficients in the tower models are consistently negative after 1930, and relatively stable. They are briefly positive around the 1929 crash. In comparison, the slope coefficients in the dividend model show relatively high volatility over time.

[INSERT FIGURE II HERE]

In accordance with prior literature, the benchmark for assessing the predictive performance of a regression model is the prevailing historical mean. This predicts the return from τ to $\tau+k$ as follows:

$$\bar{r}_{\tau,\tau+k} = \frac{1}{\tau-k}\sum_{t=1}^{\tau-k} r_{t,t+k} \,. \tag{10}$$

Separately for each of the three regression models and the two return horizons, Figure III shows how the sum of squared prediction errors of a given regression model compares to the sum of squared prediction errors of the historical mean. Specifically, for a given year T, the cumulative relative squared prediction errors are determined as follows:

$$\text{Cumulative relative SSE} = \sum_{t=t(1910)}^{T}\left(r_{t,t+k} - \bar{r}_{t,t+k}\right)^2 - \sum_{t=t(1910)}^{T}\left(r_{t,t+k} - \hat{r}_{t,t+k}\right)^2 \tag{11}$$

[INSERT FIGURE III HERE]

If the cumulative relative SSE is positive, the regression predictions perform better than does the simple prediction based on the historical mean. Starting with the overall performance from 1910-2010, it is evident from Figure III that the in-sample performance of the tower model carries over to the out-of-sample analysis. Predictive regressions using tower information lead to squared errors that are lower than those of the historical mean. In addition, the tower model is not surpassed by models that use the dividend price ratio as a single or additional variable.

These differences take some time to form. Until the 1940s, the tower model performance exhibits relatively large fluctuations, and is below benchmark for several years. Its consistency is largest for the five-year horizon. After World War II, the tower model leaves the benchmark behind. Though there are some episodes in which the relative advantage recedes, the charts do not indicate that there is a lasting shift or break in the performance. Furthermore, results of the tower model do not critically depend on whether the non-negativity constraint is imposed. Differences are so small that often they are barely visible in the figure.

Table 2 shows out-of-sample R^2 statistics, which are based on squared prediction errors as described in Section II, as well as mean absolute errors.[14] Results are presented for the 1910-2010 period and for several sub-periods. In each case, the estimation sample starts in 1871, and the Shiller data are used throughout. First of all, the statements derived from visual inspection of Figure III are supported by the statistical tests. Over the entire period as well as for sub-periods starting in 1945, 1970 or 1980, the performance of the tower model is statistically significant. Before 1945, the tower model shows reliable predictability only for the five-year return horizon. The dividend model outperforms the historical mean until 1945 if three-year or five-year horizons are considered, and underperforms afterwards. Adding the dividend price ratio to the tower information does not increase predictive power after 1945. An examination of the mean absolute error (MAE) shows that these conclusions do not change when moving from squared prediction errors to mean absolute errors.

[INSERT TABLE 2 HERE]

[14] An examination of mean prediction errors does not show any conspicuous patterns; for the tower model, they are not significantly different from zero.

The differences between the errors of unconstrained and constrained predictors are interesting because they could help to differentiate between the two possible explanations for why tower building predicts returns. The constraint is motivated by an equilibrium approach. Even if risk-aversion is very low, one would not expect risk premia to be negative since the stock market is not only risky, but also positively correlated with consumption risk. If tower building is indicative of overvaluation, by contrast, there is no reason for ruling out non-negative expected returns. For the tower models, results are mixed. Imposing the constraint tends to increase predictive performance for the one-year horizon and decrease it for the three-year horizon. In most cases, differences are relatively small. Thus, the equilibrium constraint does not help to decide between overvaluation or risk premia explanations for the predictive ability of tower information.

In order to assess the economic significance of superior predictive performance, one can examine the performance of portfolio strategies. I follow Campbell and Thompson (2008), Welch and Goyal (2008) and Rapach, Strauss and Zhou (2010) and compute average utility differences between different strategies. To facilitate the interpretation and to avoid problems from estimating the variance of multi-period returns,[15] I examine annual portfolio returns.

Assuming that an investor has mean-variance preferences of the form

$$Utility = \mathrm{E}[Return] - \frac{\gamma}{2}\mathrm{Var}[Return], \tag{12}$$

today's (today = τ) optimal equity investment w* is obtained through:

$$\max_{w} \quad Utility = R^{f}_{\tau,\tau+1} + w\hat{R}_{\tau,\tau+1} - \frac{\gamma}{2}w^2\hat{\sigma}^2_{\tau,\tau+1} \quad \Rightarrow \quad w^* = \frac{1}{\gamma}\frac{\hat{R}_{\tau,\tau+1}}{\hat{\sigma}^2_{\tau,\tau+1}} \tag{13}$$

[15] Cf. Campbell, Lo, MacKinlay (1997), ch. 2.

where R^f denotes the simple risk-free rate, $\hat{R}_{\tau,\tau+1}$ is the predicted simple equity premium, and $\hat{\sigma}^2_{\tau,\tau+1}$ is the predicted stock return variance. For the three regression models, out-of-sample forecasts of the equity premium are derived as above. To capture time variation in stock return variance, $\hat{\sigma}^2_{\tau,\tau+1}$ is estimated through the variance of the 60 monthly S&P 500 returns ending in December of year τ.[16] As in Campbell and Thompson (2008), Welch and Goyal (2008) and Rapach, Strauss and Zhou (2010), the risk-aversion parameter γ is set to 3 and portfolio weights are constrained to lie in the interval [0, 1.5] in order to rule out unrealistic strategies involving short-sales or heavy leverage. The same set of choices is made for a strategy based on the trailing historical mean. The only difference is that the expected return forecasts are replaced by the average excess return observed until year τ.

Let m_j and s_j^2 denote the sample mean and variance of the returns from a portfolio strategy based on regression forecasts, while m_0 and s_0^2 denote the sample mean and variance of a strategy based on the trailing historical mean. The utility difference obtains as

$$\Delta U = \left(m_j - \frac{\gamma}{2} s_j^2 \right) - \left(m_0 - \frac{\gamma}{2} s_0^2 \right) \tag{14}$$

Consider a situation in which the investor hires a fund manager to implement the strategy. The fund manager charges an annual management fee, determined as a percentage f of the current assets under management. This would reduce the strategy return by f, while not affecting the variance. Setting the fee f equal to the ΔU without the management fee would reduce the after-fee utility gain to zero. The utility difference in (14) can therefore be interpreted as the maximum

[16] This is the same choice as in Campbell and Thompson (2008). The variance of monthly returns is annualized through multiplication with 12. Data are obtained from Robert Shiller's website.

annual management fee that the investor would be willing to pay for the use of regression-based forecasts.

Table 3 summarizes the results. Over the 1910-2010 period, the tower model leads to a utility gain of 1.40%. The gain is negative over the years from 1910 to 1945 (-0.90%), but it quickly becomes positive. Further analysis shows that the gain is already positive (1.75%) for the 1930-1945 period. For the most recent period, the thirty years from 1980-2010, the gain is 4.25%. Adding dividend information does not greatly change the utility compared to the tower model.[17]

[INSERT TABLE 3 HERE]

The utility gains suggest that the statistical significance documented in Table 2 goes along with economic significance. For example, the post-1945 utility gains are much larger than the expense ratios incurred by mutual fund investors during that period (cf. Latzko (1999)).

Determinants of Tower Building

Two possible reasons for the predictive power of tower building are that (i) it captures credit market conditions, and therefore, risk aversion, and (ii) it proxies for market sentiment, and therefore, overvaluation. The analysis in the previous section produced evidence that is consistent with both explanations. In this section, the question shall be addressed from a different

[17] It may seem surprising that the dividend-based strategy yields positive utility gains in periods in which the dividend model does not produce superior return predictions. The picture can be explained by noting that an investor can benefit from a trading strategy through both better average performance and reduced variance (cf. Welch and Goyal, 2008).

angle. Examining whether tower building is related to credit market conditions or sentiment could shed light on the validity of the different explanations.

I use the following variables to capture credit market conditions:

- the annual change in the volume of real estate loans at all commercial banks, deflated with the US consumer price index (CPI). CPI as well as loan data from years following 1947 are from *Federal Reserve Economic Data* (Fred).[18] Loan data from years prior to 1947 are from the US All Bank Statistics.[19]

- the credit spread, defined as yield on Baa-rated bonds minus yield on Aaa-rated bonds. Data are from Fred.

To capture sentiment, I use

- the equity share in new issues. Among the components of the sentiment index of Baker and Wurgler (2006), this is the one with the longest data history. Data are from Jeff Wurgler's website.[20]

or, alternatively,

- the sentiment index of Baker and Wurgler (2006). To maximize the number of observations, I take the values of the old sentiment index prior to 1965. Data are from Jeff Wurgler's website.

Loan and spread data begin with information from the year 1919, the equity issue data begin with 1927 information, and the sentiment index is available beginning with data from 1934.

---

[18] http://research.stlouisfed.org/fred2/data/.

[19] http://fraser.stlouisfed.org/publications/allbkstat/

[20] http://pages.stern.nyu.edu/~jwurgler/

Table 4 presents regressions in which the construction of large towers as measured by the variable $LargeStart_t$ is explained through the lagged loan growth, the lagged credit spread, and the lag of one of the two sentiment variables. Coefficients are estimated with OLS, the standard errors with Newey and West (1987) and a lag length chosen according to Newey and West (1994). The regressions are highly significant, and the coefficients have the expected sign: Building activity is higher after periods of high loan growth, low credit spreads, and high sentiment. Sensitivity analyses show that the conclusions are not affected if credit spreads are defined differently (Baa yield minus long-term treasury yield, Aaa yield minus long-term treasury yield), or if contemporaneous values of the predictive variables are included in the regression. As the construction of large towers can be explained by both credit market conditions and sentiment, the data support both the rational risk aversion story and the irrational overvaluation story. In years subsequent to 1933, for which the sentiment index is available, sentiment is the dominating predictor for building activity: The adjusted R^2 in a regression with just the sentiment index is 19.5%, compared to 14.9% if the regression includes only loan growth and credit spread.

[INSERT TABLE 4 HERE]

IV. Robustness

To assess the robustness of the results from Section III, I begin by examining alternative variable definitions and predictors. Subsequently, I present results for international stock markets. The variations are tested for the three-year return horizon.

Variations of Additional Predictors

1) Instead of using the log dividend price ratio as an additional predictor, I use the log 10-year price-earnings ratio computed by Robert Shiller.

2) I replace the log dividend price ratio with the log dividend *yield*, defined as the log of dividends minus the log of *lagged* index values.

3) Rapach, Strauss, and Zhou (2010) show that a simple average of return predictions from individual regressions models produces superior out-of-sample performance. I consider the full set of 15 variables studied by Rapach, Strauss, and Zhou (2010):

   Dividend–price ratio; dividend yield; earnings–price ratio; dividend–payout ratio; historical S&P 500 volatility; book-to-market ratio of the Dow Jones Industrial Average; net NYSE equity issues scaled by NYSE market capitalization; three-month treasury bill rate; long-term government bond yield; lagged return on long-term government bonds; term spread, default yield spread (Baa minus Aaa); difference between long-term corporate bond and long-term government bond returns; inflation rate; investment-to-capital ratio (ratio of aggregate investment to aggregate capital for the entire economy.

   Rapach, Strauss, and Zhou (2010) begin their out-of-sample analysis with the year 1965, as some variables are not available from years prior to 1947. To apply their combination approach to the 1910-2010 period, I suggest the following procedure: At time *t*, use any individual prediction that is based on a regression with 10 observations or more. The number of variables that enter the prediction increases from 5 in 1910 and 10 in 1933 to 15 in 1960.[21] In Rapach, Strauss, and Zhou (2010), the arithmetic mean performs best over their full sample period. Therefore, I also take the simple average of the individual predictions.

[21] Requiring a minimum of 20 observations does not affect conclusions. The out-of-sample R^2 for the entire sample decreases to –0.065, while the 1945-2010 sub-period sees a modest improvement from 0.062 to 0.069.

4) Ferreira and Santa-Clara (2011) decompose the equity return as follows:

   (1+return) = (1 + PE-ratio growth rate) × (1 + earnings growth rate) × (1 + dividend price ratio)

   They suggest predicting the return components separately and then aggregating the partial predictions to obtain a return forecast. I implement one of their most successful specifications, in which the expected PE-ratio growth rate is set to zero, the expected earnings growth is estimated using the growth rate over the previous 20 years, and the expected dividend price ratio is estimated using the current dividend price ratio. The horizon is set to three years, which means that the earnings growth rate is taken to be the trailing average of past three-year earnings growth rates. The return forecast is converted into a forecast for the equity premium by assuming the current risk-free rate to hold over the horizon.[22]

Variations of the Tower Variable

5) In the construction of the variable *LargeStart*, public buildings such as city halls and capitols were excluded. In a variation, I include such buildings.

6) Journalists and analysts often pay special attention to towers that break a height record. To measure record-tower building, I identify the years in which construction began on a tower that would set a new US height record. The information is recorded in the following dummy variable:

   $RecordStart_t$: One if construction of a tower breaking the contemporary US record was begun in year t, zero otherwise.

[22] In a sensitivity analysis, I assume perfect foresight of risk-free rates. Conclusions are not affected.

The list of record-breaking towers is given in Table A1 in the appendix. In contrast to *LargeStart*, *RecordStart* shows little autocorrelation (the first-order autocorrelation coefficient is 0.103). To capture potential delays in the association with stock market returns, I include two further lags. In this variation, the specification is thus

$$r_{i,t,t+3} = b_0 + b_1 RecordStart_{it} + b_2 RecordStart_{i,t-1} + b_3 RecordStart_{i,t-2} + u_{i,t,t+3}$$

Variations of Return Variables

7) Instead of using the Shiller data from 1871 to 2010, I link the Shiller data from 1871 to the end of 1926 with the CRSP data from the end of 1926 to 2010.

8) Rather than using log returns for the predictive regressions and the computation of out-of-sample performance statistics, I use simple returns.

Table 5 summarizes the results of the robustness checks by presenting the out-of-sample R^2. They are shown for both the full 1871-2010 period and the 1910-1945 and 1945-2010 sub-periods.

[INSERT TABLE 5 HERE]

Over the entire sample, the price-earnings ratio leads to a better performance than does the dividend price ratio, but it does not exceed the one of the tower model. Between 1910 and 1945, the price-earnings ratio leads to a superior performance, but the performance after 1945 is disappointing. Using the dividend yield instead of the dividend price ratio reduces performance. The combination forecast of Rapach, Strauss, and Zhou (2010) makes use of a wide range of variables. It performs well after 1945, but its out-of-sample R^2 remains below that of the tower models. Before 1945, the relative performance is negative. With the sum of the parts method of

Ferreira and Santa-Clara (2011), the picture reverses. Outperformance is (marginally) significant before 1945, and then declines. Together, these findings corroborate the previous results. The performance of predictions based on tower-building activity is not only statistically significant. Over the entire sample and post-1945 they also win the race against a large number of models that have been studied intensively in the literature.

Including public buildings in the construction of the tower variables does not strongly affect the results. Replacing *LargeStart*, which is a general measure of tower-building activity, by information about construction starts of record-breaking towers improves predictive performance until 1945. After World War II, however, the picture reverses. Further analysis shows that the predictive performance of the model with the *RecordStart* variables is not significant for the most recent 1980-2010 period, for which *LargeStart* still showed a significant out-of-sample R^2 (cf. Table 2). Investors looking for predictors of future returns therefore should focus on large towers rather than on record-breaking towers.

Linking the Shiller data from 1871 to the end of 1926 with the CRSP data from the end of 1926 to 2010 does not greatly change the results relative to using the Shiller data for 1871 to 2010. In a further variant that is not reported in the table, I conduct the 1945-2010 analysis with the CRSP data only, i.e. the estimation begins with the year 1926 rather than 1871. Conclusions are not affected. Using simple rather than logarithmic returns does not lead to conspicuous changes in the results.

The predictive power of tower building activity that was documented in the previous section is therefore robust to variations in model specification.

International Evidence

Finally, I examine whether there is international evidence for the predictive power of tower building activities. I examine Datastream country stock market indices denominated in US dollars, and include all countries for which such indices are available. (The countries are listed in the Appendix.) In addition to addressing the individual countries, I also study the Datastream World ex US index. Again using the Emporis database, I define the following measure of tower-building activity on a country-by-country basis:

$LargeStart_{it}$ =

$$\ln\left(\frac{1+\text{Number of towers taller than THRESHOLD}_{\text{it}}\text{ begun in country i in year t}}{1+\text{Average annual number of towers taller than THRESHOLD}_{\text{it}}\text{ begun in country i in t-1, t-30}}\right)$$

where $\text{THRESHOLD}_{\text{it}}$ is the average height of towers of more than 50 meters, begun in years t-1,t-30 in country i. For the analysis of the world ex US index, "country i" in these definitions is replaced by "world ex US". As above, I study logarithmic excess returns over the US risk-free rate, and focus on three-year returns. They are denoted by $r_{i,t,t+3}$.

Separately for each country and the world ex US index, I run the dividend+tower model regression:

$$r_{i,t,t+3} = b_0 + b_1 dp_{it} + b_2 LargeStart_{it} + u_{i,t,t+3} \tag{15}$$

As in the previous sections, the coefficients are estimated with OLS, their standard errors according to Hodrick (1992).

I will also report mean coefficients across countries, whose precision is difficult to assess. I suggest the following simulation procedure to estimate confidence intervals under the null hypothesis that the tower variable does not have any influence on returns:

1) Randomly reshuffle the countries, i.e. for each country i, randomly draw a country without replacement from the entire set of countries, j=1,…,37.

2) For country i, determine the tower variable by using the corresponding values of country j. If country j was chosen for country i, for example, the value of $LargeStart_{it}$ is replaced by the value of $LargeStart_{jt}$ for each t.

3) Run regressions (14) with the original return and dividend variables and the reshuffled tower variables and determine the mean (across countries) of the estimated coefficients.

4) Repeat 1) to 3) 10,000 times.

In the table, I report the 99% confidence intervals of the simulated mean coefficients.

Results are presented in Table 6. An inspection of the country-by-country regressions does not reveal striking patterns related to the influence of tower building. Some coefficients are significant, but this would be expected, given the large number of regressions. The mean coefficient of tower models, however, is negative (-0.15) and outside the simulated 99% confidence interval. It is close to the coefficient on *LargeStart* when the analysis of US data is restricted to the 1972-2010 time period (-0.17, reported at the bottom of the table). Note, too, that the coefficient for the United States is insignificant even though *LargeStart* leads to superior out-of-sample forecasts in the 1970-2010 period (cf. Table 2).

[INSERT TABLE 6 HERE]

An analysis of the world ex US index corroborates the finding that *LargeStart* helps predicts future returns. The construction of large towers significantly (t-stat=-3.11) predicts lower future returns. Overall, the findings are therefore similar to the US evidence for the same time period.

V. Conclusion

In this paper, I have shown that construction starts of skyscrapers predict subsequent stock returns. The predictive ability exceeds that of alternatives such as the prevailing historical mean, predictions based on dividend price ratios and recently suggested combination forecasts (Rapach, Strauss, and Zhou, 2010; Ferreira and Santa-Clara, 2011). Predictability is mainly driven by the post-World War II period, and it does not seem to decline in more recent periods.

One explanation for the documented patterns is that tower building is indicative of over-optimism. Widespread over-optimism could lead not only to tower building, but also to overvalued stock markets. The rational asset pricing explanation is that during periods of low risk aversion, financing of large-scale projects such as skyscrapers is easier, and expected returns are lower. It is generally difficult to disentangle rational and irrational explanations for patterns in long-run returns. In this paper, I provide indirect evidence that is consistent with both views. A non-negativity constraint on return forecasts, which can be motivated by rational asset pricing theory, does not lead to an unambiguous improvement of forecasts. Furthermore, the construction of large towers is found to be correlated both with investor sentiment and with credit market conditions.

Appendix

Table A1: Towers Expected to Break the US Record at the Time When Construction Began

Name	Height	Started	Finished
Western Union Telegraph Building	70	1872	1875
New York Tribune Building	79.25	1873	1875
World Building	94.18	1889	1890
Manhattan Life Building	106	1893	1894
Park Row Building	119.18	1896	1899
Singer Building	186.57	1906	1908
Metropolitan Life Tower	213.36	1907	1909
Woolworth Building	241.4	1910	1913
Church Missionary Building	243.84	1926	
Chrysler Building	318.92	1928	1930
Empire State Building	381	1930	1931
One World Trade Center	417	1966	1972
Willis Tower	442.14	1970	1974
One World Trade Center	541.33	2006	
Chicago Spire	609.61	2007	

Table A2: Datastream Country Indices Used in the International Analysis

Country	First data in	Country	First data in
ARGENTINA	1993	MALAYSIA	1986
AUSTRALIA	1973	MEXICO	1989
AUSTRIA	1973	NETHERLANDS	1973
BELGIUM	1973	NEW ZEALAND	1988
BRAZIL	1994	NORWAY	1980
CANADA	1973	PHILIPPINES	1988
CHILE	1989	POLAND	1994
CHINA A	1994	PORTUGAL	1990
DENMARK	1973	SINGAPORE	1973
FINLAND	1988	SOUTH AFRICA	1973
FRANCE	1973	SPAIN	1987
GERMANY	1973	SWEDEN	1982
GREECE	1990	SWITZ.	1973
HONG KONG	1973	TAIWAN	1988
INDONESIA	1990	THAILAND	1987
IRELAND	1973	TURKEY	1989
ITALY	1973	UK	1969
JAPAN	1973	VENEZUELA	1990
KOREA	1987	WORLD EX US	1973

References

Ang, A., and G. Bekaert, 2007, Return predictability: Is it there? *Review of Financial Studies* 20, 651–707.

Baker, M. and J. Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645-1680.

Barr, J., 2010, Skyscraper height, *Journal of Real Estate Finance and Economics* 38, 567-597.

Boudoukh, J., M. P. Richardson, and R. F. Whitelaw, 2008, The myth of long-horizon predictability, *Review of Financial Studies* 21, 1577–605.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press, Princeton).

Campbell, J. Y., R. J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Campbell, J. Y., and S. B. Thompson, 2008, Predicting the equity premium out of sample: Can anything beat the historical average? *Review of Financial Studies* 21, 1509–31.

Clark, T. E., and K. D.West, 2007, Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics* 138, 291–311.

Cochrane, J., 2008, The dog that did not bark: A defense of return predictability, *Review of Financial Studies* 21, 1533-1575.

Condit, C.W., 1964, *The Chicago School of Architecture: A History of Commercial and Public Building in the Chicago Area, 1875-1925* (University of Chicago Press, Chicago).

Diebold, F. X., and R. S. Mariano, 1995, Comparing predictive accuracy, *Journal of Business and Economic Statistics* 13, 253–63.

Fama, E. F., and K. R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.

Ferreira, M.A., and P. Santa-Clara, 2011, Forecasting stock market returns: The sum of the parts is more than the whole, *Journal of Financial Economics* 100, 514–537.

Fisher, K.L, and M. Statman, 2006, Market timing in regressions and reality, *Journal of Financial Research* 29, 293–304.

Freeman, M.C., 2011, The time-varying equity premium and the S&P 500 in the twentieth century, *Journal of Financial Research* 34, 179–215.

Goyal, A., and I. Welch, 2003, Predicting the equity premium with dividend ratios, *Management Science* 49, 639–54.

Hodrick, R. J., 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357-386.

Landau, S.B., and C.W. Condit, 1996, *Rise of the New York Skyscraper, 1865-1913* (Yale University Press, New Haven).

Latzko, D.A., 1999, Economies of scale in mutual fund administration, *Journal of Financial Research* 22, 331-339.

Lawrence, A., 1999, The Skyscraper Index: Faulty Towers! Property Report, Dresdner Kleinwort Benson Research.

Lawrence, A., J. Hsu, W. Luo, and V. Chan, 2012, Skyscraper Index. Bubble Building. Research Report, Barclays Bank.

Lettau, M., S. Ludvigson, and J. Wachter, 2008, The declining equity premium: What role does macroeconomic risk play? *Review of Financial Studies* 21, 1653-1687.

Malkiel, B. G., 2004, Models of stock market predictability, *Journal of Financial Research* 27, 449-459.

Newey, W. K., and K. D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–8.

Newey, W. K., and K.D. West, 1994, Automatic lag selection in covariance matrix estimation, *Review of Economic Studies* 61, 631-53.

Rapach, D., J. Strauss, and G. Zhou, 2010, Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *Review of Financial Studies* 3, 821-62.

Thornton, M., 2005, Skyscrapers and business cycles, *The Quarterly Journal of Austrian Economics* 8, 51–74.

Welch, I., and A. Goyal, 2008, A Comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–508.

White, H., 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–38.

Table 1: Explaining Returns on the US Stock Market with Information Related to the Start of Large Towers

	1-year horizon (k=1)		3-year horizon (k=3)		5-year horizon (k=5)	
	1871-2010	1926-2010	1871-2010	1926-2010	1871-2010	1926-2010
Panel A: Dividend yield as predictor – Dividend model						
dp_t	0.046	0.082	0.168	0.267*	0.295	0.415*
	(1.21)	(1.51)	(1.45)	(1.74)	(1.59)	(1.74)
Adj. R^2	0.004	0.018	0.043	0.102	0.088	0.189
N	139	83	137	81	135	79
Panel B: Tower building as predictor – Tower model						
$LargeStart_t$	-0.042***	-0.055***	-0.112***	-0.151***	-0.173**	-0.200***
	(-2.68)	(-2.77)	(-2.36)	(-2.50)	(-2.33)	(-2.51)
Adj. R^2	0.051	0.084	0.135	0.246	0.207	0.321
N	139	83	137	81	135	79
Panel C: Dividend yield and tower building as predictors – Dividend+tower model						
dp_t	0.020	0.027	0.096	0.117	0.198	0.241
	(0.53)	(0.46)	(0.84)	(0.68)	(1.05)	(1.00)
$LargeStart_t$	-0.040***	-0.050**	-0.101**	-0.132*	-0.153**	-0.164**
	(-2.44)	(-2.17)	(-2.14)	(-1.96)	(-2.05)	(-2.09)
Adj. R^2	0.046	0.076	0.144	0.254	0.242	0.371
N	139	83	137	81	135	79

Note: The log excess return on the US stock market from t to t+k is regressed on the log dividend price ratio (dp) and a variable containing information about the building starts of large towers in the United States. $LargeStart_t$ relates the number of towers that were larger than a trailing average and that were started in year t to the number of such starts in the 30 years before t. Data from 1871 to 2010 are constructed by Robert Shiller based on the S&P 500 and other series; Data from 1926 to 2010 are for the CRSP value-weighted market portfolio. T-statistics are based on White (1980) for the one-year horizon, and on Hodrick (1992) for horizons longer than one year. Coefficients of regression constants are not reported.

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

Table 2: Analysis of Out-of-Sample Prediction Errors

			Dividend model		Tower model		Dividend+Tower model	
		Hist. mean	unconstr.	constr.	unconstr.	constr.	unconstr.	constr.
One-year return horizon								
1910-2010	R^2_{os}	-	-0.003	-0.003	0.041**	0.046**	0.024**	0.028**
	MAE	0.194	0.194	0.194	0.190	0.189	0.191	0.191
1910-1945	R^2_{os}	-	-0.010	-0.010	-0.017	-0.010	-0.037	-0.033
	MAE	0.228	0.229	0.229	0.230	0.229	0.232	0.232
1945-2010	R^2_{os}	-	0.005	0.005	0.097***	0.101***	0.082***	0.087***
	MAE	0.173	0.173	0.173	0.164	0.164	0.166	0.165
1970-2010	R^2_{os}	-	-0.007	-0.007	0.074**	0.079**	0.056*	0.064*
	MAE	0.183	0.184	0.184	0.176	0.176	0.178	0.177
1980-2010	R^2_{os}	-	-0.041	-0.041	0.090**	0.097**	0.062*	0.073*
	MAE	0.179	0.183	0.183	0.171	0.170	0.174	0.173
Three-year return horizon								
1910-2010	R^2_{os}	-	0.011	0.021*	0.104***	0.100***	0.078***	0.083***
	MAE	0.339	0.337	0.335	0.321	0.321	0.325	0.324
1910-1945	R^2_{os}	-	0.059***	0.060***	-0.025	-0.025	0.008	0.008
	MAE	0.408	0.396	0.395	0.413	0.413	0.406	0.406
1945-2010	R^2_{os}	-	-0.045	-0.025	0.245***	0.236***	0.161***	0.170***
	MAE	0.286	0.292	0.289	0.249	0.250	0.262	0.260
1970-2010	R^2_{os}	-	-0.074	-0.043	0.243***	0.229***	0.161**	0.152**
	MAE	0.292	0.302	0.298	0.254	0.256	0.267	0.269
1980-2010	R^2_{os}	-	-0.227	-0.185	0.294**	0.276**	0.119*	0.106
	MAE	0.293	0.324	0.319	0.246	0.249	0.275	0.277
Five-year return horizon								
1910-2010	R^2_{os}	-	0.027**	0.073**	0.217***	0.217***	0.191***	0.234***
	MAE	0.431	0.425	0.415	0.381	0.381	0.388	0.377
1910-1945	R^2_{os}	-	0.186***	0.191***	0.115**	0.115**	0.239***	0.239***
	MAE	0.487	0.439	0.438	0.458	0.458	0.425	0.425
1945-2010	R^2_{os}	-	-0.141	-0.054	0.327***	0.327***	0.149***	0.231**
	MAE	0.390	0.417	0.401	0.320	0.320	0.360	0.342
1970-2010	R^2_{os}	-	-0.208	-0.116	0.262**	0.262**	0.157**	0.171*
	MAE	0.366	0.402	0.387	0.314	0.314	0.336	0.333
1980-2010	R^2_{os}	-	-0.556	-0.416	0.257**	0.257**	-0.036	-0.004
	MAE	0.359	0.447	0.427	0.309	0.309	0.365	0.359

Note: Out-of-sample forecasts of three-year stock returns are generated using (i) a model including only the dividend price ratio (dividend model); (ii) model based on the building counts of large towers (tower model); (iii) a model combining the two predictors (dividend+tower model). Constrained forecasts are forced to be non-negative. Stock market data are constructed by Robert Shiller based on the S&P 500 and other series. Statistical significance of the R^2_{os} statistic (out-of-sample R^2 relative to the historical mean) is based on the Clark and West (2007) MSPE-adjusted statistic, computed with Newey and West (1987) standard errors. MAE is mean absolute error.
***Significant at the 1% level.
**Significant at the 5% level.
*Significant at the 10% level.

Table 3: Utility Gains of Regression-Based Investment Strategies

	Dividend model	Tower model	Dividend+tower model
1910-2010	0.58	1.40	1.27
1910-1945	-0.56	-0.90	-1.00
1945-2010	1.16	2.66	2.51
1970-2010	1.90	4.38	4.32
1980-2010	0.68	4.25	4.18

Note: Table entries are the estimated annual fees (in percent) that mean-variance investors with a risk aversion of three would be willing to pay for using regression forecasts of the equity premium rather than the trailing historical mean. The results obtain from an out-of-sample analysis in which optimized portfolios using (i) a model including only the dividend price ratio; (ii) a model based on the building counts of large towers; or (iii) a model combining the two predictors are compared with optimized portfolios based on the trailing mean. Portfolio weights are restricted to line in the interval [0, 1.5].

Table 4: Determinants of Tower Building

	Depvar: *LargeStart*$_t$				
Sample:	1919-2010	1927-2010	1934-2010	1934-2010	1934-2010
Real estate loan growth (t-1,t-2)	5.790***	5.630**	4.369*	5.860**	
	(2.90)	(2.27)	(1.65)	(2.08)	
Baa-Aaa spread (t-1)	-25.045	-29.376*	-22.988	-35.287	
	(-1.41)	(-1.76)	(-0.75)	(-0.99)	
Equity share in new issues (t-1)		2.477			
		(1.63)			
Sentiment index (t-1)			0.379***		0.483***
			(3.62)		(3.54)
constant	-5.932***	2.477	0.379***	-6.044**	-0.158
	(-2.82)	(-2.40)	(-1.64)	(-2.03)	(-0.62)
p(regression)	.005***	.016**	.000***	.055*	.000***
Adj. R^2	0.159	0.205	0.254	0.149	0.195
N	90	82	75	75	75

Note: The table shows whether loan growth, credit spreads, the equity share in new issues, and the sentiment index from Baker and Wurgler (2006) explain the building starts of towers in the United States. *LargeStart*$_t$ relates the number of towers that were larger than a trailing average and that were started in year t to the number of such starts in the 30 years before t. Coefficients are estimated with a linear regression. T-statistics (in parentheses) are based on Newey and West (1987) with an automated lag length selection. Different sample periods and specifications are motivated by data availability.

***Significant at the 1% level.
**Significant at the 5% level.
*Significant at the 10% level.

Table 5: Out-of-Sample R²s from Sensitivity Analyses

Variation relative to base case	Evaluation period	Non-tower predictor only		Tower predictor only	
		unconstrained	constrained	unconstrained	constrained
Base case (Table 2)					
	1910-2010	0.011	0.021*	0.104***	0.100***
	1910-1945	0.059***	0.060***	-0.025	-0.025
	1945-2010	-0.045	-0.025	0.245***	0.236***
Variations of non-tower predictors					
1) PE ratio instead of dp ratio	1910-2010	0.088**	0.094***		
	1910-1945	0.186***	0.177***		
	1945-2010	-0.024*	0.001*		
2) Dividend yield instead of dp ratio	1910-2010	-0.010	0.011		
	1910-1945	0.042	0.043		
	1945-2010	-0.068	-0.025		
3) Combination forecast as in RSZ	1910-2010	0.010	0.010		
	1910-1945	-0.037	-0.037		
	1945-2010	0.062*	0.062*		
4) Sum of the parts forecast as in FSC	1910-2010	0.027*	0.028*		
	1910-1945	0.044*	0.045*		
	1945-2010	0.010	0.010		
Variations of tower predictors					
5) Public buildings included	1910-2010			0.104***	0.101***
	1910-1945			-0.023	-0.023
	1945-2010			0.243***	0.236***
6) Three lags of record-tower building dummy	1910-2010			0.147***	0.147***
	1910-1945			0.097**	0.098**
	1945-2010			0.205***	0.202***
Variatons of return definitions					
7) Shiller data until 1926, then CRSP	1910-2010	-0.011	-0.002	0.106***	0.103***
	1910-1945	0.047**	0.055**	0.021*	0.021*
	1945-2010	-0.091	-0.082	0.211***	0.205***
8) Discrete returns instead of log returns	1910-2010	0.014	0.018	0.097***	0.097***
	1910-1945	0.062**	0.062**	-0.061	-0.061
	1945-2010	-0.033	-0.027	0.249***	0.249***

Note: This table presents out-of-sample R² statistics (R^2_{os}) relative to the historical mean for several variations relative to the base case in Table 2. Constrained forecasts are forced to be non-negative. A three-year horizon is employed. Combination forecasts are obtained as in Rapach, Strauss, and Zhou (RSZ) (2010), while sum of the part forecasts follow Ferreira and Santa-Clara (FSC) (2011); the performance of these forecasts is evaluated directly without using them as a predictor in a regression. Statistical significance of the R^2_{os} statistic is based on the Clark and West (2007) MSPE-adjusted statistic, computed with Newey and West (1987) standard errors. Cells are left empty if the variation does not change the specification relative to the base case.

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

Table 6: Forecasting International Stock Market Returns

	dp_{it}	$LargeStart_{it}$	Adj. R^2	N
Argentina	-0.33 (-2.29)**	-0.11 (-0.47)	0.14	15
Australia	0.59 (1.46)	-0.19 (-1.19)	0.05	36
Austria	0.76 (1.83)*	-0.17 (-1.31)	0.05	36
Belgium	0.21 (0.94)	-0.18 (-1.13)	0.01	36
Brazil	0.26 (0.87)	0.25 (0.87)	0.12	14
Canada	-0.17 (-0.56)	-0.03 (-0.23)	0.00	36
Chile	0.83 (2.92)***	0.17 (0.86)	0.32	19
China	-0.88 (-2.06)**	0.43 (0.96)	0.34	14
Denmark	0.16 (0.81)	0.17 (0.43)	-0.03	36
Finland	0.13 (0.55)	-1.04 (-2.38)***	0.11	20
France	0.29 (0.73)	-0.08 (-0.61)	0.01	36
Germany	0.15 (0.54)	-0.01 (-0.04)	-0.04	36
Greece	1.25 (3.30)***	-2.56 (-0.69)	0.51	18
Hongkong	1.08 (3.77)***	-0.02 (-0.16)	0.38	36
Indonesia	0.33 (1.23)	-0.26 (-0.58)	0.07	18
Ireland	0.37 (1.18)	-0.20 (-0.37)	0.06	36
Italy	0.56 (1.93)*	-0.18 (-1.33)	0.02	36
Japan	0.48 (1.90)*	-0.02 (-0.23)	0.15	36
Malaysia	1.09 (1.70)*	0.25 (0.88)	0.41	22
Mexico	1.03 (2.09)**	0.43 (1.40)	0.29	19
Netherlands	0.43 (1.68)*	0.06 (0.52)	0.08	36
New Zealand	0.16 (0.24)	-0.22 (-0.88)	-0.04	20
Norway	0.86 (2.48)***	0.19 (0.61)	0.19	28
Philippines	0.58 (1.40)	-0.48 (-1.95)*	0.25	20
Poland	0.01 (0.03)	0.47 (1.23)	0.08	14
Portugal	0.69 (2.19)**	-0.28 (-0.79)	0.27	18
Singapore	0.75 (2.84)***	-0.13 (-1.04)	0.28	36
South Africa	0.53 (1.37)	-0.17 (-0.78)	0.11	36
South Korea	1.12 (2.06)**	0.16 (0.43)	0.22	21
Spain	0.35 (0.84)	0.07 (0.45)	-0.05	21
Sweden	0.37 (0.92)	-0.31 (-1.15)	0.08	26
Switzerland	0.03 (0.11)	-0.17 (-0.65)	-0.05	36
Taiwan	0.48 (1.46)	0.15 (0.62)	0.13	20
Thailand	0.59 (1.68)*	-0.70 (-1.68)*	0.28	21
Turkey	0.28 (0.49)	-0.40 (-1.05)	0.05	19
United Kingdom	0.56 (1.36)	-0.06 (-0.53)	0.17	39
Venezuela	0.16 (0.74)	-0.30 (-0.58)	0.04	18
Mean coefficient across countries with simulated 99% range for mean under H_0: no relationship for towers				
	0.44	-0.15***		
	[0.403, 0.455]	[-0.069, 0.057]		
Regressions for world ex US index with world ex US tower data as well as for US restricted to (1973-2010)				
World ex US	0.42 (1.53)	-0.49 (-3.11)***	0.22	36
US (1973-2010)	0.12 (0.67)	-0.17 (-1.61)	0.33	36

Note: The three-year log excess return of country i is regressed on the country's log dividend price ratio (dp_{it}) and information about the building starts of towers in country i. $LargeStart_{it}$ relates the number of towers that were larger than a country trailing average and that were started in year t, to the number of such starts in the 30 years before t. Return data are from Datastream, and tower data are from Emporis. T-statistics are based on Hodrick (1992). The simulated range for mean coefficients is from a simulation in which values for the tower-building related variables are randomly reshuffled across countries. Coefficients of regression constants are not reported.

***Significant at the 1% level.

**Significant at the 5% level.

*Significant at the 10% level.

Figure I: Large Tower Building Activity in the US over Time. The figure shows the variable *LargeStart*$_t$, defined as

$$\ln\left(\frac{1+\text{Number of towers with height above THRESHOLD\_b}_t\ \textit{begun}\ \text{in year t}}{1+\text{Average annual number of tower starts above THRESHOLD\_b over years t-1, t-30}}\right)$$

where THRESHOLD_b$_t$ is the average height of towers of over 50 meters that were begun in years t-1, t-30. Data are from Emporis.

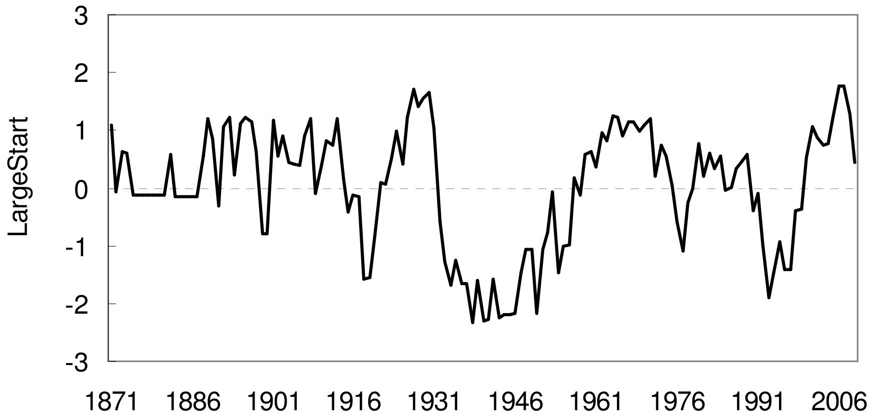


Figure II: Recursive Coefficient Updates. The figure shows the slope coefficients estimates from recursive regressions for two different models and three return horizons (k = 1, 3 or 5 years):

Dividend model: $r_{t,t+k} = b_0 + b_1 dp_t + u_{t,t+k}, \quad t = 1,...,\tau - k$

Tower model: $r_{t,t+k} = b_0 + b_1 LargeStart_t + u_{t,t+k}, \quad t = 1,...,\tau - k$

where $r_{t,t+k}$ is the log excess return on the US stock market from year t to t+k; dp is the log dividend price ratio; LargeStart_t relates the number of towers that were larger than a trailing average and that were started in year t to the number of such starts in the 30 years before t. Data from 1871 to 2010 are constructed by Robert Shiller based on the S&P 500 and other series.

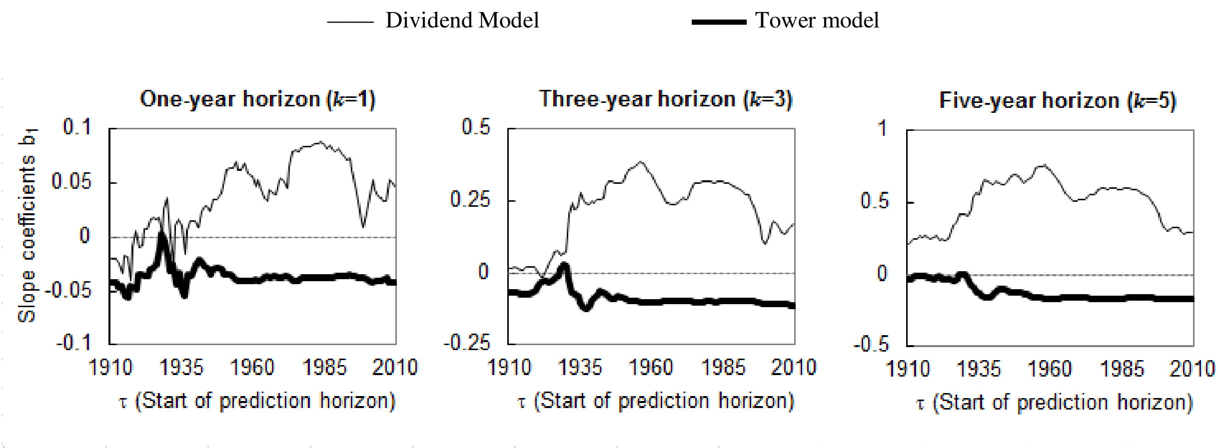
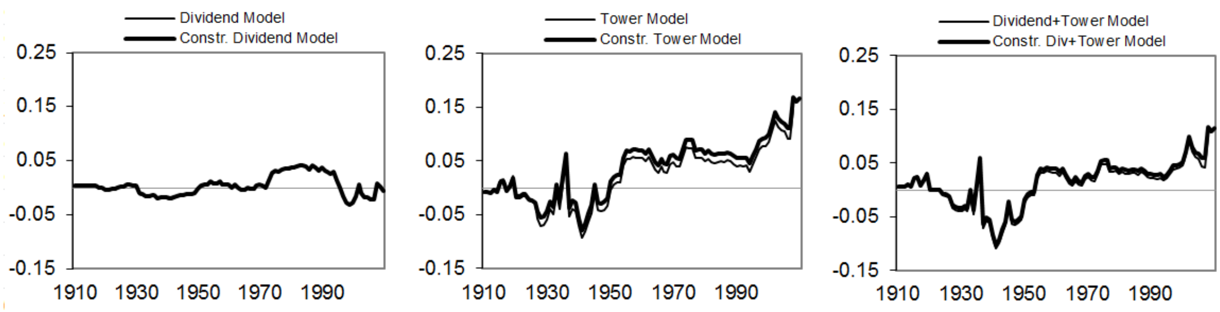


Figure III: Cumulative Out-of-Sample Performance Relative to the Prevailing Mean. Out-of-sample forecasts of k-year stock returns are generated using (i) a model including only the dividend price ratio (dividend model); (ii) model based on the building counts of large towers (tower model); (iii) a model combining the two predictors (dividend+tower model). Constrained forecasts are forced to be non-negative. Predictions made in t are denoted by $\hat{r}_{t,t+k}$. The figure plots the relative sum of squared predictions errors

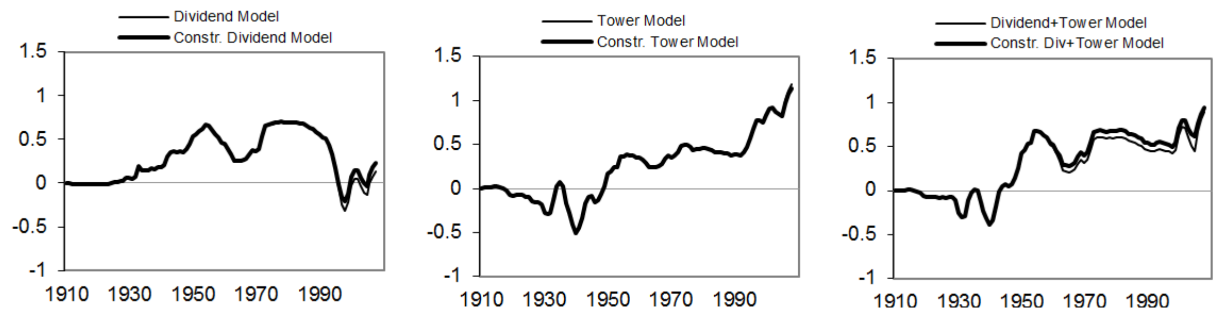
$$\text{Cumulative relative SSE } = \sum_{t=t(1910)}^{T}\left(r_{t,t+k} - \bar{r}_{t,t+k}\right)^2 - \sum_{t=t(1910)}^{T}\left(r_{t,t+k} - \hat{r}_{t,t+k}\right)^2$$

where the historical k-year mean return over years 1871 to t is denoted by $\bar{r}_{t,t+k}$. A positive cumulative relative SSE shows that a regression model performed better than the historical mean.

Cumulative relative SSE for the one-year horizon (k = 1)



Cumulative relative SSE for the three-year horizon (k = 3)



Cumulative relative SSE for the five-year horizon (k = 5)