Chapter 5

Leverage and Stock Return Predictability: Evidence from U.S. Panel Data

5.1 Introduction

The predictability of stock returns is a widely studied topic; with this contribution, we revisit the question and establish, both theoretically and empirically, the significant forecast power of firm leverage (asset-to-equity ratios) in the U.S. stock market. In accordance with the efficient market hypothesis, we establish a simple prediction model, which also aligns with the traditional CAPM and capital structure theory. In this model, firm leverage is a potential predictor of returns on equity at the firm level. In a marked departure from most existing studies, we test the prediction model using panel data, instead of an integrated market index, from the U.S. stock market and thus can sufficiently use leverage information included in the cross-section. The empirical results confirm the good forecasting performance of the proposed model for long-run equity return predictions.

The predictability of stock returns has attracted attention from both practitioners and academics. The earliest studies mainly focused on testing the efficient market hypothesis (Balvers et al., 1990; Basu, 1977; Malkiel, 2003). In addition, various financial variables have been considered, with multiple eco-

1This chapter is based on Qiao (2016).
nometric tools, to forecast individual or aggregate stock returns and test for 
weak-form stock market efficiency, but the results are not consistent across dif-
ferent markets or periods; see, for example, Beechey et al. (2000). Moreover, 
various economic and firm-level data have supported tests of the semi-strong-
form and strong-form efficient market hypothesis. But, similar to investiga-
tions of the weak-form efficient market hypothesis, those studies often offer 
contradictory results (Givoly & Lakonishok, 1979; Groenewold & Kang, 1993; 
Summers, 1986). In light of these conflicting findings, some studies point out 
that the presence of stock return predictability does not necessarily contradict 
the efficient market hypothesis, because the risk premium and systemic risk 
of a stock or portfolio could be time-varying, due to the evolution of business 
cycles or changes in the economic environment. Estimating the time-varying 
risk premium is a difficult task though. Many scholars have proposed variables 
to capture the risk premium, such as the term structure slope, stock volatility, 
or investor sentiment (Baker & Wurgler, 2006; Campbell, 1987; French et al., 
1987b). The dividend-price ratio plays a special role in the study of stock return 
predictability. The Campbell-Shiller linearization of the return identity implies 
that, in absence of rational price bubbles, the dividend-price ratio should pre-
dict either stock returns or dividend growth. Previous studies attempt to jointly 
test the predictability of stock returns and dividend growth; see, for example, 
Cochrane (2007).

Among these many studies of stock return predictability, two shortcomings 
are evident. First, most studies of return predictability focus on a single stock 
market index or portfolio, likely due to the econometric difficulty of jointly test-
ing return predictability for various stocks or portfolios. Investigating a single 
time-series may fail to account for useful information present in the cross-
section. However, the meaning of the presence of return predictability in a 
panel data setting remains uncertain. The panel data equivalent of stock re-
turn predictability arguably might imply that at least one stock or portfolio in 
the sample exhibits evidence of a time-varying conditional expected return; see 
Dockery and Kavussanos (1996). Such a claim requires a joint test, but no clear 
consensus exists about which test might be most appropriate. Second, many 
studies lack a theoretical framework and seem driven solely by empirical ins-
ights. In a common procedure, studies suggest a possible leading variable for 
stock returns based on some intuition, then test whether this variable actually 
has predictive power. From this perspective, evidence of stock return predict-
ability is not incompatible with classic finance theory, such as the CAPM or the Modigliani-Miller theory.

Even if they offer some intuition or compelling stories to illustrate why such variables might forecast returns, they rarely offer a rigorous, theoretical framework to derive the prediction equation, nor do they offer clear hypotheses about which values would constitute a reasonable magnitude for the parameters.

We try to overcome both shortcomings. Specifically, we suggest a predictor of expected stock returns with a strong theoretical foundation: financial leverage. On the one hand, the traditional Capital Asset Pricing Model implies that the expected excess return is the product of the risk premium and asset’s Beta. On the other hand, the Modigliani-Miller theorem states that a firm’s value is independent from its capital structure but firms’ financial leverage ratios might depend on business-cycle. Thus, the evolution of expected excess equity returns may be driven mainly by variation in two variables, the risk premium and firms’ or portfolios’ leverage ratios, as measured by the total asset-to-equity ratio. Because variation in the risk premium is much smaller and slower than changes in the leverage ratio, we ignore time-variation in the risk premium and focus on the predictive power of leverage ratios for the equity premium. In this simple environment, the statistical point becomes more transparent; specifications that also account for changes in the risk premium can only increase the evidence of return predictability. Furthermore, using yearly U.S. data from 1951 to 2013, we construct a balanced panel of 25 portfolios, which offers two key advantages. Econometric tests often require a balanced data set, and adopting fewer portfolios rather than individual firm data circumvents the statistical issues associated with large N and small T. Moreover, we construct the portfolios such that each portfolio has the same unlevered beta, based on first-step contemporaneous regressions of the returns on aggregate stock market returns. In turn, we establish clear theoretical priors for the estimated effect of the lagged leverage ratio on the portfolio returns, namely, each portfolio should have the same estimated coefficient, and the magnitude of the coefficient should be roughly equal to the average historical risk premium. These clear theoretical priors also circumvent many econometric issues associated with more agnostic hypotheses in joint tests of return predictability in panel data settings. For example, we could relax the requirement that all coefficients should be the same for each portfolio, but it would become much harder to develop a joint test in such an environment. Formally, return predictability in this relaxed setting
indicates that "existing assets in the sample exhibit time varying expected return," but no consensus exists regarding how such a test of joint significance should best be formulated. Instead, restricting the coefficients to be of the same magnitude fits within the standard panel data framework.

We adopt conventional fixed/random effect panel regressions to test the joint significance of the predictive power of the leverage ratio. They show that the leverage ratio is a significant predictor of stock returns; moreover, the estimated coefficient is close to its theoretical value.

The rest of this chapter is organized as follows: In the next section, we describe the data set. Then we introduce our motivation and derive our model specification. Subsequently, we describe why and how we construct 25 portfolios, which allow us to conduct a convenient standard panel data regression to test our hypothesis. Finally, we present our empirical results and conclusions.

5.2 Data Description

We use various data sources to construct yearly individual firms’ and aggregate market returns, as well as financial leverage ratios. To construct a proxy for the firm-level leverage ratios, we extracted data on total assets and total debt from COMPUSTAT and Moody’s Industrial, Public Utility Transportation, Bank & Finance Manuals. The equity returns and number of shares outstanding came from CRSP. Aggregate stock market returns and interest rates reflect the Fama-French and liquidity factors, available in WRDS. Finally, the Flow of Funds (FoF) account of the Federal Reserve provides a proxy for the total assets and debt of the aggregate market. We construct yearly unbalanced panel data for 7563 U.S. firms for the period from 1952 to 2013.

We define the unlevered returns, or asset returns, of individual firms and The aggregate market as

$$R_{i,t+1}^A = \frac{E_{i,t}}{E_{i,t} + B_{i,t}} R_{i,t+1}^E + \frac{B_{i,t}}{E_{i,t} + B_{i,t}} R_{i,t+1}^B,$$

where total equity and equity returns are denoted by $E_{i,t}$ and $R_{i,t}^E$, and total debt and debt returns are denoted by $B_{i,t}$ and $R_{i,t}^B$. This definition implies that we assume firms’ total assets are financed by either equity or debt. We use lagged

---

2 Approaches for joint significance tests have been suggested in prior literature (Hjalmarsson, 2006, 2010; Westerlund et al., 2015)
values of $E_{i,t}$ and $B_{i,t}$, known at time $t$, to transform the levered returns to unlevered returns. Defining the leverage ratio as $L_{i,t} = \frac{E_{i,t} + B_{i,t}}{E_{i,t}}$, we can rewrite equation (5.1) as

$$R_{i,t+1}^A = R_{i,t+1}^E \frac{1}{L_{i,t}} + R_{i,t+1}^B (1 - \frac{1}{L_{i,t}}).$$  \hspace{1cm} (5.2)

Furthermore, while ignoring default risk, we assume the return on debt is approximately equal to the risk-free interest rate $r_t$, which yields an easy-to-implement expression of the unlevered excess return,

$$R_{i,t+1}^u = \frac{R_{i,t+1}}{L_{i,t}}. \hspace{1cm} (5.3)$$

where $R_{i,t+1}^u$ is unlevered excess return $R_{i,t+1}^A - r_{t+1}$, and $R_{i,t+1}$ is the equity excess return, $R_{i,t+1}^E - r_{t+1}$. We present the summary statistics in Table 5.1.

Table 5.1. Means and Standard Deviations for Key Variables, U.S., 1952 - 2013

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Market</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market Excess Return (%)</td>
<td>61</td>
<td>7.74</td>
<td>18.61</td>
</tr>
<tr>
<td>Market Leverage</td>
<td>61</td>
<td>1.74</td>
<td>0.30</td>
</tr>
<tr>
<td>Market Unlevered Excess Return (%)</td>
<td>61</td>
<td>4.26</td>
<td>11.16</td>
</tr>
<tr>
<td>Individual Firms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excess Return (%)</td>
<td>106605</td>
<td>11.68</td>
<td>53.02</td>
</tr>
<tr>
<td>Leverage</td>
<td>123271</td>
<td>6.13</td>
<td>122.42</td>
</tr>
<tr>
<td>Unlevered Excess Return (%)</td>
<td>104344</td>
<td>6.38</td>
<td>34.98</td>
</tr>
<tr>
<td>Total Equity, Market Value (Billion $)</td>
<td>123271</td>
<td>1.72</td>
<td>10.07</td>
</tr>
<tr>
<td>Total Assets, Market Value (Billion $)</td>
<td>123271</td>
<td>3.01</td>
<td>15.70</td>
</tr>
</tbody>
</table>

The yearly excess returns of the market portfolio and the risk-free rate are taken from the Fama-French and liquidity factors in the Wharton Research Data Services. Market value of the total equity of the aggregate market comes from the FoF account of the Federal Reserve. We obtained firm-level data from COMPUSTAT North-America, Fundamentals Quarterly and Fundamentals Yearly; missing values of book equity are replaced by data from Davis, Fama, and French (2000). Excess returns are taken from CRSP. We define leverage as the market value of total assets/market value of equity, for both the aggregated market and individual firms. Unlevered excess returns are calculated as excess returns$(1 + 1)/leverage(t)$. The data reflect an unbalanced panel of 7563 firms for 61 years.
5.3 Model

To establish an unlevered, or asset, version of the unconditional CAPM, we assume the market portfolio of the asset is mean-variance efficient, such that

\[ R_{i,t+1}^u = \beta_i \lambda_{t+1} + e_{i,t+1}. \quad (5.4) \]

This assumption could explain some well-known phenomena, such as the leverage effect. Considering the relation between unlevered and levered excess returns, we rewrite the definition of asset returns from Equation (5.3), equivalently as

\[ R_{i,t+1} = L_{i,t} \beta_i \lambda_{t+1} + L_{i,t} e_{i,t+1}. \quad (5.5) \]

Eq. (5.5) can be viewed as a variant of the Unlevered CAPM developed in Chapter 2. In this chapter, I deleverage the asset return by moving the asset-to-equity ratio to the right hand side of Eq. (2.3).

Prior research has documented that asset betas are more stable over time than equity betas. Moreover, from a theoretical perspective, the unlevered asset betas, which reflect systematic firm risk, should not be affected by financing choices. We therefore postulate that expected unlevered asset returns are equal to the asset beta multiplied by the risk premium,

\[ E_t(R_{i,t+1}^u) = \beta_i \lambda_{t+1}. \quad (5.6) \]

Substituting the expression for expected asset returns in Eq. (5.5) gives

\[ E_t(R_{i,t+1}) = L_{i,t} \beta_i \lambda_{t+1}. \quad (5.7) \]

This equation can be estimated using standard methods, because the leverage ratio \( L_{i,t} \) is observable at time \( t \). We ignore time variation of the risk premium \( \lambda_{t+1} \) and conjecture that the expected excess equity return of a firm should be proportional to its lagged leverage ratio \( L_{i,t} \). Moreover, the size of the coefficient equals the asset beta, multiplied by the average risk premium. We test this conjecture in the following section.
5.4 Portfolio Construction

Several challenges arise when testing for return predictability at the individual firm level in a panel data setting, using Eq. (5.7). First, individual firm data are highly unbalanced, creating various pragmatic and econometric issues. Second, the heterogeneity in unlevered betas prevents estimates of the model with a standard fixed or random-effect panel data regression, because the coefficients in Eq. (5.7) are firm specific, whereas standard methods estimate a single coefficient for the slope. Third, our sample includes 61 years in time-series and 7653 firms in cross-section, so \( N \) is much larger than \( T \).

To circumvent these issues, we propose to group the firms into 25 portfolios and calculate value-weighted excess equity returns and leverage ratios, in a distinct way. In particular, we construct a balanced panel of 25 portfolios, and sort in such a way that the unlevered Beta, \( \beta_i \), is the roughly the same for each of the resulting portfolios and close to one. The advantage of this approach is that we can estimate a common slope for each of the portfolios, essentially restricting the betas to be the same for every portfolio in the estimation. Moreover, this approach implies that the estimated coefficient should be equal to the average risk premium.

Therefore, we first estimate a contemporaneous unlevered CAPM for each individual firm and sort them asset betas, from small to large. Next, we group firm \( i \) into the \( n_i \)th portfolio, according to the following heuristic:

\[
n_i = \text{mod}(k_i, 25) \text{ if } \text{mod}(k_i, 25) \neq 0, \text{ otherwise } n_i = 25,
\]

such that \( k_i \) is the rank of firm \( i \). This procedure selects firms in systematic, unbiased order, such that the asset betas of 25 portfolios are approximately equal.

In the next step, we calculate the value-weighted average equity excess returns and unlevered returns for each of the 25 portfolios. For the average unlevered excess return, we use total assets as weights, whereas for the average equity excess return, we employ total equity as weights. We also calculate total equity and total assets for the 25 portfolios by summing the values of all their individual firms and, as for the individual firms, define portfolios’ leverage as the asset-to-equity ratio. With this approach, we can construct a balanced panel data set of 25 cross-sections and 61 years. Table 5.2 contains the descriptive statistics for the key variables across the 25 portfolios.
Table 5.2. Means and Standard Deviations for Key Variables of 25 Portfolios, 1952 - 2013

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>S.D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess Return (%)</td>
<td>1525</td>
<td>8.73</td>
<td>19.34</td>
</tr>
<tr>
<td>Leverage</td>
<td>1525</td>
<td>1.79</td>
<td>0.43</td>
</tr>
<tr>
<td>Unlevered Excess Return (%)</td>
<td>1525</td>
<td>4.78</td>
<td>11.83</td>
</tr>
<tr>
<td>Total Equity, Market Value (Billion $)</td>
<td>1525</td>
<td>137.00</td>
<td>187.00</td>
</tr>
<tr>
<td>Total Assets, Market Value (Billion $)</td>
<td>1525</td>
<td>239.00</td>
<td>310.00</td>
</tr>
</tbody>
</table>

We construct 25 portfolios from an unbalanced panel of 7583 firms for 61 years. The firm-level data come from COMPSTAT North-America, Fundamentals Quarterly and Fundamentals Yearly, and the missing values of book equity are replaced by data from Davis, Fama, and French (2000). Excess returns are taken from CRSP. We define leverage as the market value of total assets/market value of equity, for both the aggregated market and individual firms. Unlevered excess returns are calculated as excess returns(t + 1)/leverage(t). We estimate a contemporaneous CAPM for each individual firm and sort the firms by asset betas, from small to large. Next, we group firm i into the n'th portfolio, according to the following heuristic:

\[ n^i = \text{mod}(k^i, 25) \text{ if } \text{mod}(k^i, 25) \neq 0, \text{ otherwise } n^i = 25, \]

where \( k^i \) is the rank of firm i.

5.5 Result

5.5.1 Testing for Return Predictability with Panel Data Regressions

Before applying standard fixed or random effects panel data regressions to our portfolios, we need to write the unlevered CAPM as the following expression,

\[ R_{i,t+1} = L_{i,t} \beta \lambda_{t+1} + e_{i,t+1}, \]  

(5.9)

and our hypothesis is that

\[ \begin{align*}
\beta_1 = \beta_2 = \ldots = \beta_{25} &= 1, \\
\lambda_1 = \lambda_2 = \ldots = \lambda_T &= \lambda.
\end{align*} \]  

(5.10)

Because of our elaborated construction of 25 portfolios, it is reasonable to assert that they have identical betas. Therefore, if our data do not include sample selection bias, none of the betas should be different from the market portfolio’s
Leverage and Stock Return Predictability: Evidence from U.S. Panel Data 111

beta, which equals one. In this case, the model can be simplified to

\[ R_{i,t+1} = \lambda L_{i,t} + \varepsilon'_{i,t+1}, \] (5.11)

which is a standard balanced panel model with homogeneous slopes, \( \lambda \). Consequently, we can use both fixed and random effect panel regressions to estimate the model,

\[ R_{i,t+1} = \alpha + \lambda L_{i,t} + \varepsilon'_{i,t+1}. \] (5.12)

The results are in Table 5.3.

Table 5.3. Fixed and Random Effect Panel Data Regression for 25 Portfolios

<table>
<thead>
<tr>
<th>Model</th>
<th>( \alpha )</th>
<th>( \lambda )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Effect Panel Data Regression</td>
<td>-0.07***</td>
<td>0.09***</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(-3.10)</td>
<td>(6.88)</td>
<td></td>
</tr>
<tr>
<td>Random Effect Panel Data Regression</td>
<td>-0.03</td>
<td>0.06***</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(-1.22)</td>
<td>(5.58)</td>
<td></td>
</tr>
<tr>
<td>Prais-Winsten Panel Data Regression</td>
<td>-0.01</td>
<td>0.05*</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(-0.11)</td>
<td>(1.82)</td>
<td></td>
</tr>
</tbody>
</table>

This table reports standard fixed and random effect panel data regressions for the following return prediction model:

\[ R_{i,t+1} = \alpha + \lambda L_{i,t} + \varepsilon'_{i,t+1}. \]

The testing assets are 25 portfolios, constructed from an unbalanced panel of 7563 firms for 61 years. Firm-level data come from COMPSTAT North-America, Fundamentals Quarterly and Fundamentals Yearly, and the missing values of book equity are replaced by data from Davis, Fama, and French (2000). Excess returns are taken from CRSP. We define leverage as the market value of total assets/market value of equity, for both the aggregated market and individual firms. Unlevered excess returns are calculated as excess returns\( (t+1)/\text{leverage}(t) \). We estimate a contemporaneous CAPM for each individual firm and sort the firms by asset betas, from small to large. Next, we group firm i into the nth portfolio, according to the following heuristic:

\[ n' = \text{mod}(k', 25) \text{ if mod}(k', 25) \neq 0, \text{ otherwise } n' = 25, \]

where \( k' \) is the rank of firm i. The t-statistics for the fixed effect regression and z-statistics for both random effects and Prais-Winsten regressions are reported in parentheses. *, **, and *** denote significance at 10%, 5%, and 1%, respectively.

As expected, both the fixed and random effects panel regressions exhibit positive, significant \( \lambda \), equal to 0.09 and 0.06, respectively. The magnitude is slightly greater than the estimated risk premium, at 0.04. However, in the random effects regression, the Wald-test shows that we cannot reject the claim that
the slope is significantly different from 0.04. These results support our assertion that the lagged leverage ratio can predict equity returns and be consistent with theoretical models. In turn, our findings provide evidence that the lagged leverage ratio can effectively predict stock returns.

5.5.2 Robustness

Traditional fixed or random effect panel regressions account for cross-sectional correlation and time-series auto-correlation. However, both correlations are common in financial markets, especially cross-sectional correlations (Frees, 1995; Kolari & Pynnönen, 2010; Roll & Ross, 1994). As a robustness check, we therefore implement Prais-Winsten regressions to capture potential correlations in both the cross-section and time-series. The time-series auto-correlation is modeled by a AR(1) model; the cross-sectional correlation is estimated by correlated panel corrected standard errors: see Baltagi et al. (2007). The last two rows of Table (5.3) contain the results.

In Prais-Winsten Panel Data Regression, we find a more reasonable estimate of the slope, equal to 0.05, which is closer to the historical average risk premium, at 0.04, that is positively significant at the 10% level. The estimated constant is equal to -0.01, that is, much smaller than the fixed and random effects panel regression results of -0.74 and 0.03, respectively. Therefore, it appears more consistent with the theoretical model when we consider both cross-section correlation and time-series auto-correlation. Overall, this evidence offers more support for our conjecture: The lagged leverage ratio is an effective leading variable to predict stock returns.

5.6 Out of Sample Prediction

The preceding evidence showing that the leverage ratio is an efficient predictor of equity returns derives from a panel data regression of the whole sample, meaning that the findings mainly reflect in-sample performance. As an additional check, in this section, we implement an out-of-sample prediction to test the model’s forecasting performance. We estimate the parameters using a training sample from 1953 to 1999, which generates forecasted values for the equity excess returns in the testing sample, which runs from 2000 to 2013. That is, using the training sample, we run the panel regression in Eq. (5.12), and obtain estimations of the parameters, \( \hat{\alpha} \) and \( \hat{\lambda} \). For the testing sample, we then
Leverage and Stock Return Predictability: Evidence from U.S. Panel Data 113

generate forecast values of equity excess returns,

\[ \hat{R}_{i,t} = \hat{\alpha} + \hat{\lambda}L_{i,t-1}, \]

(5.13)

where i spans from 1 to 25, and t covers 2000 to 2013. Furthermore, we calculate squared forecasting errors for all firms in each year,

\[ \hat{e}_{i,t} = (\hat{R}_{i,t} - R_{i,t})^2. \]

(5.14)

Relative to leverage-based forecasting, the benchmark model relies on simple average forecasting. Specifically, we calculate the average equity excess return for each firm in the training sample,

\[ \bar{R}_i = \frac{\sum_{t=1953}^{1999} R_{i,t}}{1999 - 1953 + 1}. \]

(5.15)

Using \( \bar{R}_i \) as a forecast in our testing sample, we recalculate the squared forecasting errors as

\[ \bar{e}_{i,t} = (\bar{R}_i - R_{i,t})^2. \]

(5.16)

We present the summary statistics for of \( \hat{e}_{i,t} \) and \( \bar{e}_{i,t} \) in Panel A of Table 5.4.

Generally, the out-of-sample performance of leverage-based forecasting is better than simple average forecasting. The mean of the squared forecasting errors for leverage-based forecasting is 3.74% (its squared-root is about 19.34%), smaller than the simple average forecasting error of 3.99% (its squared-root is about 19.97%). Yet the standard deviations of leverage-based forecasting and simple average forecasting are 5.67% and 6.05%, respectively, so leverage-based forecasting is more stable. To test the significance of the difference, we turn to a matched-pairs t-test and Wilcoxon signed-ranks test for the squared forecasting errors of the two kinds of forecasting. Panels B and C of Table 5.4 contain the results.

In Table 5.4, both the matched-pairs t-test and Wilcoxon signed-ranks test indicate significant differences in the forecasting errors. The average difference between leverage-based forecasting and simple average forecasting’s squared forecasting error is -0.25, which is negative and significant at the 1% level. In another words, the leverage prediction model’s forecasting error is significantly smaller than that of the simple average. To control for disnormality in forecasting error, we also implement a Wilcoxon signed-ranks test. The observed pos-
Table 5.4. Our-of-Sample Squared Forecasting Errors for 25 Portfolios

<table>
<thead>
<tr>
<th>Panel A: Summary Statistics</th>
<th>Obs.</th>
<th>Mean (%)</th>
<th>S.D. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leverage Ratio Prediction ($\hat{\epsilon}$)</td>
<td>350</td>
<td>3.74</td>
<td>5.67</td>
</tr>
<tr>
<td>Simple Average Prediction ($\bar{\epsilon}$)</td>
<td>350</td>
<td>3.99</td>
<td>6.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Matched-pairs T-test</th>
<th>Mean (%)</th>
<th>S.D.(%)</th>
<th>T-Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diff.=$\hat{\epsilon} - \bar{\epsilon}$</td>
<td>-0.25***</td>
<td>1.23</td>
<td>3.86</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Wilcoxon Signed-ranks Test</th>
<th>Observed</th>
<th>Expected</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>151</td>
<td>175</td>
<td>0.99</td>
</tr>
<tr>
<td>Negative</td>
<td>199***</td>
<td>175</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 3.1 reports the summary statistics, Matched-pairs T-test and Wilcoxon Signed-ranks Test of squared forecasting errors of leverage prediction model and simple average prediction in our testing sample, from 2000 to 2013. The squared forecasting errors of leverage prediction model is calculated as $\hat{\epsilon}_{it} = (\hat{R}_{it} - R_{it})^2$, and $\hat{R}_{it} = \hat{\alpha} + \hat{\lambda} L_{it-1}$ is the fitted value of equity excess return of firm $i$ at year $t$. Using our training sample, from 1953 to 1999, we obtain the estimation $\hat{\alpha}$ and $\hat{\lambda}$ by a panel data regression $R_{it+1} = \alpha + \lambda L_{it} + \epsilon_{it+1}$. Additionally, $\bar{\epsilon}$ is the squared forecasting errors of a simply average prediction, namely $\bar{\epsilon}_{it} = (\bar{R}_{it} - R_{it})^2$. We calculate the average equity excess return of each firm in our training sample and use the average as a forecast, $\bar{R}_{it}$. The 25 testing portfolios are constructed from an unbalanced panel of 7563 firms. Firm-level data come from COMPUSTAT North-America, Fundamentals Quarterly and Fundamentals Yearly, and the missing values of book equity are replaced by data from Davis, Fama, and French (2000). Excess returns are taken from CRSP. We define leverage as the market value of total assets/market value of equity, for both the aggregated market and individual firms. Unlevered excess returns are calculated as excess return of the 1st leveraged portfolio. We estimate a contemporaneous CAPM for each of the individual firms and sort the firms by asset betas, from small to large. Next, we group firm $i$ into the $n'$th portfolio, according to the following heuristic:

$$n' = \text{mod}(k', 25) \text{ if } \text{mod}(k', 25) \neq 0, \text{ otherwise } n' = 25,$$

where $k'$ is the rank of firm $i$. *, ** and *** denote the significance at 10%, 5% and 1% respectively.

Positive difference is just 151, smaller than the expected value of 175. Yet the observed negative difference is 199, significantly larger than the expected value of 175 at the 1% level. In summary, out-of-sample forecasting results confirm that the leverage ratio can improve predictions of equity excess returns. Leverage-based forecasting’s error is smaller and more stable than that of simple average forecasting.
5.7 Conclusion

This chapter proposes a theoretical return-forecasting model that can integrate the traditional CAPM and capital structure theory. Many previous studies lack a theoretical framework and seem driven solely by empirical insights. In addition, most existing studies of return predictability focus on a single stock market index or portfolio, likely due to the econometric difficulty of jointly testing return predictability for various stocks or portfolios. Our return-forecasting model and the proposed novel portfolio construction approach can overcome the two shortcomings. Empirical study implies that portfolios’ leverage forecasts portfolio returns. A one percentage-point change in the leverage ratio yields a 0.04 percentage-point change in the expected excess return in the subsequent year. To circumvent econometric issues associated with potential heterogeneity in the slope coefficients of the cross-section, we construct 25 portfolios with identical asset betas, using yearly U.S. data from 1951 to 2013. With these portfolios, we also avoid technical problems associated with an unbalanced panel. This process may cause the loss of some information when we construct the portfolios, but reintegrating that information would only make our results stronger.

Using conventional fixed and random effects panel data regressions, we estimate and test the significance of the leverage ratio’s coefficient. Theoretically, the coefficient should be positively significant and close to the risk premium of 0.04, if our data are free of sample selection bias. Both of the fixed and random effects panel regressions indicate a significant, positive slope of the leverage ratio, and the estimations are roughly the same order of magnitude as the risk premium.

As a robustness check, we consider correlations in cross-section and autocorrelation in time-series. We estimate the panel using the Prais-Winsten regression to capture both kinds of correlation. This regression reveals a significant positive slope of leverage in the results, and its magnitude of 0.05 is quite close to the risk premium. An insignificant negative estimation of the constant also is consistent with our theoretical model.

The out-of-sample test provides further evidence for our conclusion. We estimate our model’s parameters using a training sample and implement the forecast with the testing sample. In a comparison with the simple average forecast, we find that both the matched-pairs t-test and Wilcoxon signed-ranks test reveal that leverage-based forecasting exhibits smaller, more stable forecasting
errors in the testing sample. This conclusion is consistent with our previous findings.

Unlike most prior research, our proposed method tests return predictability using panel data instead of an individual stock or market index. This approach enables us to extract more useful information from the data set. Therefore, our method and results address some gaps in previous works. In general, we find that the leverage ratio is an important leading variable of stock returns, as revealed by the significant positive relation between the lagged leverage ratio and current stock return.

One remaining question is the predictive power of the time-varying market risk premium. Previous studies have shown that market risk premium is lower during expansion period than during recession period. Thus, the movement of market risk premium may help to improve the performance of our prediction model. We will leave this question open for future research.