Understanding the Behavior of Distressed Stocks

Yasser Boualam*
Joao F. Gomes†
Colin Ward‡

December 21, 2012

Preliminary and Incomplete

Abstract

This paper shows how non-linearities in returns induced by delisting events can affect the inference about the behavior of delisting stocks. Because these events are both extreme and introduce a floor on expected stock returns, the correct factor model is likely to be quite non-linear. As a result the estimated alphas and loadings in standard linear models are biased. We show that although these biases can be significant for abnormal excess returns they are generally quite small for factor loadings. Empirically this is because delisting events are largely uncorrelated with systematic risk factors. After we correct these biases we see little evidence of underperformance for portfolios of distressed stocks.

*The Wharton School, University of Pennsylvania, boualam@wharton.upenn.edu
†The Wharton School, University of Pennsylvania, gomesj@wharton.upenn.edu
‡The Wharton School, University of Pennsylvania, wardcol@wharton.upenn.edu

1 Introduction

Understanding the behavior of distress stocks has proved something of a challenge for financial economists. Earlier thought, going back to at least Fama and French (1992), suggested that financial distress could perhaps be the source of the higher expected returns of value stocks. Unfortunately however, most of the recent research seems to conclude that portfolios of highly distress stocks tend to severely underperform other stocks.[1] At least as surprising is the fact that the estimated loadings of distress portfolios on standard risk factors, especially on size, are often quite large, rendering the puzzle even deeper.

This paper re-examines both the evidence and the methodology behind these studies. Our starting point is the observation that delisting is an extreme event and one that imposes a lower bound on the expected (and realized) stock return of the firm. We believe that this is unlikely to be well captured with the simple linear factor model used in most other studies.

In the next section we show theoretically that accounting for the probability of this extreme event can dramatically change the properties of expected returns and, specifically, impact our estimates of any abnormal excess returns for portfolios of highly distressed stocks, where this probability can be non-trivial. Our theory also implies there are potentially important biases in the estimated factor loadings.

We then use a Monte-Carlo simulation in Section 4 to illustrate these points quantitatively. Using empirically plausible default probabilities we show how fitting a simple linear model leads to biased coefficient estimates, and more specifically, to generally negative portfolio alphas. This is true even though our true underlying data generating process is assumed to have zero excess returns.

More constructively, these two theoretical sections also suggest a relatively simple procedure to adjust portfolio returns and estimate the correct excess returns and factor loadings. Section 5 implements this suggested correction and compares the results with those from the linear models that are standard in the literature. Our findings confirm

[1] Some examples include Dichev, (1998), Griffin and Lemmon (2002), Campbell et al (2008) and Garlappi and Yan (2011). Vassalou and Xing (2004) however reaches the opposite conclusion.

that the correct excess returns are indeed much smaller than previously estimated and in many cases not statistically different from zero. In particular, it is no longer true that the portfolios of the most highly distressed stocks exhibit strongly negative alphas.

Although we find our evidence compelling, there are a few subtle, but potentially important, issues that we do not entirely resolve in this paper. Perhaps the most significant has to do with the fact that detecting financial distress is inherently difficult. It is common in the literature to identify a "distress" or a "default" event, with stock delistings for performance related reasons, and we will follow this practice here too. As a result, we will use the terms distress, default, and delisting more or less interchangeably, although the latter is the more accurate one.

Practically, this means that we will identify highly distressed stocks as those with a very high probability of being delisted for performance related reasons. Estimating this probability accurately, then, becomes a significant step in any study of financial distress. We focus on two alternative measures of these probabilities. The first one updates the reduced form logit approach introduced by Shumway (2001) and Campbell et al (2008). For contrast, we also use a more structural measure, building on Merton's Distance to Default insights. These two alternative approaches are discussed in Sections 3 and 6, respectively.

Another possible complication has to do with the identification of the exposure of different stocks and portfolios to distress risk. The importance of separating this exposure into systematic and idiosyncratic components is rarely acknowledged explicitly in the literature. Nevertheless, this is also paramount to any discussion on the impact of distress on expected equity returns. In the most extreme case, when defaults are entirely driven by firm-level shocks, they are of little concern for investors and we should not expect to see any sort of premium or discount on highly distressed stocks.

Section 2 shows formally how allowing for some correlation between default probabilities and the systematic risk factors introduces an additional source of bias in estimated factor loadings. Interestingly, we can show that the magnitude of this bias in factor loadings is decreasing with the size of this correlation. Section 3 then sheds light on the empirical magnitude of this source of bias by explicitly identifying and separating idiosyncratic and systematic components of distress.

We now turn to discuss the details of our calculations.

2 The Model with Delisting Returns

We begin by describing our model for the stochastic process for expected stock returns when delisting events occur and proceed to discuss the theoretical biases in trying to fit linear factor models to this process. We show that the likely magnitude of these biases depends on the both the size of the delisting probabilities and their correlation with the systematic risk factors. We estimate these in Section 3 before using Monte-Carlo methods to attempt to gauge the magnitudes of these biases in Section 4.

2.1 The Factor Model

Our starting point is the observation that the limited liability of equity investors means that bankruptcy and stock delistings effectively impose a lower bound on the value of both actual and expected stock returns. As a result it seems sensible to assume that the true process for expected excess stock returns on stock i in some arbitrary portfolio p at time t, denoted $r_{ip}(t)$ is given by the expression:[2]

$$r_{ip}(t) = \begin{cases} \tilde{r}_{ip}(t) = \alpha_p + \beta_p F(t) + \epsilon_i(t) & \text{with prob} \quad 1 - p_i(t) \\ -\delta & \text{with prob} \quad p_i(t) \end{cases} \tag{1}$$

where $p_i(t)$ is the probability that firm i will delist its stock.

The upper branch of the stochastic process (1), denoted $\tilde{r}_{ip}(t)$, describes the familiar multi-factor linear representation of expected equity returns. Here $F(t)$ is a vector of priced factors, β_p is a vector of (portfolio) factor loadings and $\epsilon_i(t)$ is idiosyncratic noise. For most practical applications we can think of $F(t)$ as including the popular Fama and French (1993) factors although in our empirical analysis we also discuss the role of momentum factors. If our factor specification is correct, the excess returns α_p will be uniformly 0. Henceforth we will assume this to be the case.

[2]Although equation (1) is quite general it is common in this literature to work with excess returns over the market portfolio and not the risk free rate.

The novelty here is the introduction of the lower branch in equation (1). This captures the fact that if the stock is delisted, which occurs with time varying probability $p_i(t)$, the expected return is the delisting return $-\delta$. Many theoretical models of endogenous default effectively imply that $\delta = 1$ although in practice this is probably a worst case scenario.[3]

2.2 Systematic Components of Distress

Our second modification to the standard factor model is the explicit discussion of both systematic and idiosyncratic components to stock delistings. Although this aspect is occasionally implicit in the literature on financial distress, we are not aware of any extant attempt to formalize its impact.

By assumption, systematic risk in our model is captured by the vector of factors $F(t)$. Thus, to the extent that default probabilities have systematic components, it seems natural to assume that the probability of default for firm i at time t obeys the following relation:

$$p_i(t) = \eta_{0,i} + \eta_{f,i} F(t) + \xi_i \tag{2}$$

This representation explicitly decomposes the probability of delisting, $p_i(t)$, into its idiosyncratic and systematic components. Here $\eta_{0,i}$ and ξ_i capture the predictable and stochastic components, respectively, of firm specific determinants of distress, while $\eta_{f,i}$ summarizes the covariance with systematic variables. Hence, if distress or delisting probabilities are entirely firm specific we expect that $\eta_{f,i} = 0$. As we show below, the (plausible) correlation between factor returns and default probabilities can potentially lead to biased empirical estimates of the factor loadings.[4]

2.3 Theoretical Biases

Taken together, the two stochastic processes (1) and (2) pose a number of challenges to many empirical studies on the effects of distress on equity returns. In effect the typical procedure for most of the existing literature is as follows:

[3] Some classical examples include Merton (1974) and Leland (1994).

[4] In our empirical implementation we will also allow for a more general non-linear correlation between $p_i(t)$ and $F(t)$.

- Estimate default probabilities $\hat{p}_i(t)$, usually using firm level data.
- Create portfolios of firms sorted by default probabilities.
- Estimate the linear factor model.

$$r_{ip}(t) = a_p + b_p F(t) + \epsilon_i(t) \tag{3}$$

and compute average excess returns, $\hat{a}_p$ and factor loadings, $\hat{b}_p$, across portfolios.

Unfortunately, if our more general model is true there are two specific problems with this simple approach:

1. The process for expected returns is not linear, and specifically is truncated from below.
2. The truncation probabilities are endogenous and correlated with return factors.

Specifically, we will now show that the standard approach generally leads to biased estimates for both alphas ($a's$) and factor betas ($b's$).

To see this observe that the correct specification for excess returns for stock i in equation (1) can be re-written as:

$$r_i(t) = (1 - p_i(t))(\beta_i F(t) + \epsilon_i(t)) + p_i(t)(-\delta) \tag{4}$$

Combining this expression with that for (2) it follows that we can write the 'true' excess return process as:

$$\begin{aligned} \mathrm{E}r_i(t) &= \mathrm{E}[(1 - (\eta_{i,0} + \eta_{f,i}F(t) + \xi_i(t)))(\beta_i F(t) + \epsilon_i(t)) - (\eta_{i,0} + \eta_{f,i}F(t) + \xi_i(t))\delta] \\ &= -\eta_{i,0}\delta + ((1 - \eta_{i,0})\beta_i - \eta_{f,i}\delta)F(t) - \sum_{f=1}^{n_f}\sum_{k=1}^{n_f} \eta_f \beta_k F_f(t) F_k(t) \end{aligned} \tag{5}$$

where the last line assumes that the shocks ϵ_i and ξ_i are idiosyncratic and uncorrelated with each other. When the (n_f) risk factors are also uncorrelated the last term is also 0 and the "reduced" form loadings are linked to the "structural" loadings by the equation:[5]

$$b_i = (1 - \eta_{i,0})\beta_i - \eta_{f,i}\delta \tag{6}$$

[5] Even if factors are somewhat correlated the final term in equation (5) is of second order importance.

Equation (6) states that we will have two types of potential biases in the linear factor estimates of returns on distressed portfolios. The first is a standard bias associated with return truncation, which implies that the linear estimates $\hat{b}$ will always be biased towards zero. The second, and possibly offsetting, bias has to do with the fact that distress is correlated with risk factors. If, plausibly, $\eta_{f,i} > 0$ so that distress probabilities load positively on risk factors, this source of bias will push the estimated linear loadings $\hat{b}$ below the true factor β's.

Equation (6) also shows that a linear reduced form regression will generally deliver downward biased estimates of the intercept α. In particular, when the true excess return is 0 this equation implies a *negative* estimated α's, which are equal to:

$$a_i = -\eta_{i,0}\delta \tag{7}$$

In this light the well documented underperformance of stocks with a high probability of delisting may not be very surprising. The question however is how large this bias can be in practice and how much of this perceived "underperformance" survives after we correct it. We think both of these findings are potentially of great importance for the literature on the performance of distressed stocks.

Intuitively equations (6) and (7) suggest a potentially important modifier to Shumway's (1997) well-known argument about the role of delisting biases. The addition of delisting returns on the CRSP sample is clearly important. However, the introduction of these sharply non-linear events should be modeled with great care. Continuing to assume that the process for expected stock returns continues to obey a simple linear factor model creates large potential biases in parameter estimates. In particular, this approach quite possibly generates sizable, and spurious, negative alphas in high distress portfolios.

The likely magnitude of these biases depends on the empirical properties of the stochastic process for the delisting probabilities (2). We study these probabilities in the next section and then use them to construct a Monte-Carlo simulation in Section 4 that attempts to quantify the sources of bias in estimated excess returns and factor loadings of linear factor models.

3 Estimating Default Probabilities

A crucial ingredient in any study of distress or default risk is a measure of the probability of delisting events, $p_i(t)$. In what follows we identify a default event with a stock delisting for performance-related reasons. Specifically we use the following delisting flags from the CRSP monthly file: 500, 550, 552, 560, 561, 574, 580, and 584. Appendix A discusses these classifications and their properties at length. For the sample period model discussed in this section this classification yields 5,994 delistings out of almost 200,000 firm-year observations.

3.1 Data Overview

Our data covers the period 1950 to 2011, although most of the analysis focuses on the period from 1970 on. Firm level data comes from combining quarterly accounting data from COMPUSTAT with monthly and daily data from CRSP. When quarterly accounting data is not available, we use annual data.

We use all industrial, standard format, consolidated accounts of USA headquartered firms in COMPUSTAT. We follow Campbell et al. (2008) and align each company's fiscal year with that of the calendar year, and then lag the accounting data by two months. Our measure of book equity follows Davis, Fama, and French (2000). From the CRSP monthly and daily file we use all stocks in NYSE, AMEX, and NASDAQ. The S&P500 index comes from the annual MSP500 file and data on the Fama and French size and book to market factors come from Ken French's website. Details about the data and our approach to construct the key variables are included in Appendix B. Table I reports the summary statistics for the variables used in our regressions.

3.2 Logistic Regressions

We forecast delisting events using two alternative approaches. For our baseline analysis we use an updated version of the reduced form logistic model proposed by Campbell et al (2008). In section 6 we also report the results when we use an implementation of Estimated Default Frequencies (EDF), proposed by Merton (1974). As we will see both of these measures offer very good forecasts of delisting events over this period, at least at the portfolio level. This is particularly true for our benchmark reduced form logit model, which is discussed here.

Formally we use maximum likelihood methods to estimate a logistic function on eight explanatory variables in a pooled estimation across all firm-years. Our methodology here differs somewhat from that of Campbell et al. (2008). They use monthly regressions and focus on predicting the probability of defaulting 12 months ahead, *conditional* on no default occurring in the 11^{th} month. Instead, we use annual rolling logit regressions that can be interpreted as estimating the probability of defaulting, at any time *within the next year*, given the information available at the beginning of the year. More precisely, we estimate these rolling regressions on an annual basis starting in December 1970 up to December 2011 to avoid any look-ahead bias.

Formally we define $p_i(t) = 1/(1 + \exp(-y_i(t)))$, where $y_i(t)$ can be approximated by the following empirical specification:

$$\begin{aligned} y_i(t) = \quad & \gamma_0 + \gamma_{EXRETAVG} EXRETAVG_i(t) + \gamma_{SIGMA} SIGMA_i(t) \\ & + \gamma_{PRICE} PRICE_i(t) + \gamma_{NIMTAAVG} NIMTAAVG_i(t) + \gamma_{TLMTA} TLMTA_i(t) \\ & + \gamma_{CASHMTA} CASHMTA_i(t) + \gamma_{RSIZE} RSIZE_i(t) + \gamma_{MB} MB_i(t) \end{aligned} \tag{8}$$

where $EXRETAVG_i(t)$ is a measure of average excess returns over the S&P500 index, $SIGMA_i(t)$ is the volatility of equity returns, $MB_i(t)$ is the market to book ratio, $NIMTAAVG_i(t)$ is a measure of profitability, $TLMTA_i(t)$ is a measure of firm leverage, $CASHMTA_i(t)$ is a measure of cash holdings, $RSIZE_i(t)$ is the relative size of the firm, and $PRICE_i(t)$ is the log stock price, capped at \$15.

The full sample logistic regression results are shown in Table II. They do not differ materially from those in Campbell et al (2008). The pseudo R-squared for these firm level estimates is nearly 40% and all of these financial and accounting ratios are immensely significant.

3.3 Probability Portfolios

Based on the estimated probabilities $\hat{p}_i(t)$ each firm is then ranked and assigned a percentile on a scale of zero to one-hundred in this empirical distribution. Next we form $p = 1, ..9$ portfolios every year in December and each firm is placed in the correct percentile portfolio. These portfolios are ranked in a symmetric and increasing order (1 through 9) as follows:[6]

- Portfolio 1: Percentiles between 0% and 5%
- Portfolio 2: Percentiles between 5% and 10%
- Portfolio 3: Percentiles between 10% and 20%
- Portfolio 4: Percentiles between 20% and 40%
- Portfolio 5: Percentiles between 40% and 60%
- Portfolio 6: Percentiles between 60% and 80%
- Portfolio 7: Percentiles between 80% and 90%
- Portfolio 8: Percentiles between 90% and 95%
- Portfolio 9: Percentiles between 95% and 100%

Although the portfolio composition is fixed over the course of a calendar year, both the probabilities and the value-weights on each stock are allowed to fluctuate over the year with the change in each firm's accounting variables and returns, respectively.

We use value weights to construct portfolio returns and incorporate the delisting returns into our portfolio return calculations we also follow Campbell et al. (2008) and simply use the CRSP delisting return when available or the lagged monthly returns otherwise.

Average delisting probabilities for each portfolio are computed using equal weights. Formally, the portfolio's December-to-December equal-weighted (or average) predicted probability for portfolio p equals:

$$\bar{p}_p(t) = \sum \hat{p}_{ip}(t)/N_p(t) \tag{9}$$

where $N_p(t)$ is the number of stocks in portfolio p at time t.

[6] As usual there is a degree or arbitrariness about these classifications. In practice nearly all delistings come from the stocks ex-ante classified in the percentiles 80-100 so the breakdowns for the first 5 or 6 portfolios are not particularly important. It is sometimes useful to create finer portfolios for the upper percentiles but there is also a concern that the number of firms in each of them will become quite low, particularly as so many are then delisted over the calendar year.

Table III documents the basic patterns of delisting probabilities, stock returns and other characteristics across our nine delisting portfolios. Average delisting probabilities, $\bar{p}_p(t)$ are quite low for the first four or five portfolios. Excess returns (over the market) are generally negative for the portfolios with a high probability of delisting. Return volatility and skewness is also much higher for these stocks. The sharp increase in return skewness is consistent with our view of delistings as highly non-linear events. As documented extensively distress portfolios are also generally made of small and low book-to-market firms.

Figure 2 depicts the time series for the average predicted probabilities of delistings, $\bar{p}_p(t)$ for each portfolio p. As can be seen these probabilities exhibit significant time variation and are noticeably higher around market downturns, suggesting that there is a possibly important systematic component to delistings, at least for some portfolios.

3.4 Actual and Predicted Delistings

Before proceeding it is instructive to investigate the accuracy of our estimated average delisting probabilities. To this effect we also construct a yearly time series of actual annual delisting events and compute the ex-post delisting *frequencies* for each portfolio. This time series is shown in Figure 3 and, visually, it seems to accord remarkably well with that in Figure 2. Both figures show that delistings are almost entirely concentrated in the top 3, or perhaps 4, portfolios.

More formally, Table IV reports the results of regressing the actual ex-post delisting frequencies on our average predicted probabilities, $\hat{p}_p$. Although the quality of fit might seem poor for the first 3 portfolios it should be noted that there is virtually no variation in the dependent variable (delistings) here. By contrast for the last 3 or 4 portfolios where default is concentrated the fit seems much more accurate. For all but the first three portfolios the estimated coefficients are also very close to 1 as we would expect if the fit is accurate.[7]

[7]No intercept is included in these regressions. This assumes that if we predict a zero probability of delisting (which never occurs), then we are forcing the actual delisting probability to be zero. Including an intercept reduces the highest R-squared to 0.66.

3.5 Systematic Components of Delisting Probabilities

Section 2 shows that when delisting probabilities have a systematic component there is a potential source of bias in our estimates of the factor loadings. To investigate this issue we now regress the average predicted portfolio delisting probabilities, $\bar{p}_p(t)$, on the systematic risk factors. Although our baseline representation in equation (2) was linear we now allow for a more general third-order polynomial expansion. Intuitively the introduction of square terms may be useful to capture periods of increased volatility, which coincide with high default probabilities. Allowing for cubic terms can capture skewness in portfolio probabilities.

To implement this equation empirically we first convert our logit's annual predicted probabilities to monthly probabilities using the following formula:

$$p_i^{mth}(t) = 1 - (1 - p_i^{ann}(t))^{1/12} \tag{10}$$

We then estimate the following empirical regression for the average delisting probabilities across portfolios on a third degree polynomial on the four Carhart factors:

$$\begin{aligned}\hat{p}_p(t) = \eta_{0,i} \quad &+ \quad \eta_{1,i} MKT(t) + \eta_{2,i} MKT^2(t) + \eta_{3,i} MKT^3(t) \\ &+ \quad \eta_{b1,i} HML(t) + \eta_{b2,i} HML^2(t) + \eta_{b3,i} HML^3(t) \\ &+ \quad \eta_{s1,i} SMB(t) + \eta_{s2,i} SMB^2(t) + \eta_{s3,i} SMB^3(t) \\ &+ \quad \eta_{m1,i} MOM(t) + \eta_{m2,i} MOM^2(t) + \eta_{m3,i} MOM^3(t) \\ &+ \quad \eta_{bs,i} SMB \times HML(t) + \eta_{bm,i} HML \times MOM(t) + \eta_{sm,i} SMB \times MOM(t) \\ &+ \quad \xi_i \end{aligned} \tag{11}$$

Table V reports both our coefficient estimates and the implied R^2 for each of the nine stock portfolios. Somewhat curiously, in light of the suggestive evidence on the time variation of delisting probabilities presented earlier, there is remarkably little evidence of covariance between delisting probabilities and risk factors. In fact almost all of this covariance comes from exposure to higher order terms in the momentum factor, and they tend to be sizable only for the last three portfolios.

These findings have two important implications. First, from a practical standpoint, it suggests that this covariance between factors and probabilities is unlikely to produce a significant bias in our estimates of factor loadings - at least not when using these baseline logistic default probabilities.

More broadly however, this evidence of little actual systematic risk in delistings indicates that equity investors should not care very much about these events. As a result we should not expect to see much risk compensation for them. With this interpretation any empirical findings of negative excess returns in high distress portfolios must be the result of some form of mis-pricing.[8]

In the next two sections we investigate the empirical implications of these findings. We proceed in two complementary steps. First we combine these estimated probabilities with numerical simulation methods to attempt to quantify the likely bias in empirical estimates. We also propose a return correction that accounts for the non-linear role of delistings on portfolio returns. Finally, in Section 5 we use this correction to provide more accurate estimates of factor loadings and excess returns across distress portfolios.

4 Numerical Simulation

The analysis in the previous two sections raises two questions:

1. Are the theoretical biases in factor loadings and excess returns quantitatively significant?; and
2. How do we obtain more accurate estimates of these coefficients?

We now tackle both of them.

4.1 Implementation

Suppose that we have a cross-section of firms that is made $p = 1, 2..9$ portfolios each made of 250 individual stocks. Each portfolio is ranked in increasing order of default probabilities, $\bar{p}_p(t)$. For simplicity we assume that each firm i in portfolio p has an equal ex-ante default probability, that is equal to the average probability for the entire

[8] Of course it is also possible, although unlikely, that we left out an important risk factor from this regression.

portfolio. Formally then $p_i(t) = \bar{p}_p(t)$. Once a stock is delisted, it is excluded from its portfolio for the rest of the year. We assume that each delisted stock is only replaced by a new stock at the beginning of the *following* year.

Assuming firm level stock excess returns follow the stochastic process (1), we can generate an artificial panel of 504 months of excess stock returns for each stock i in portfolio p by drawing realizations from the process:

$$r_{ip}(t) = \begin{cases} \beta_p F(t) + \epsilon_{ip}(t) & \text{if } \gamma_p(t) > \bar{p}_p(t) \\ -\delta & \text{else.} \end{cases} \tag{12}$$

where $\gamma_p(t) \sim U[0,1]$.

By construction there are no abnormal excess returns to these returns. The exact values of the true factor loadings, β_p are not particularly important. For consistency, however we assume that they are also equal to their empirical counterparts (reported below in section 5).

The law of large numbers implies that the (equally-weighted) average excess returns of stocks in each portfolio p are given by:[9]

$$r_p(t) = \sum_i r_{ip}(t) \simeq (1 - \bar{p}_p(t))(\beta_p F(t)) - \bar{p}_p(t)\delta \tag{13}$$

We now use our estimates of the properties of the delisting probabilities $\bar{p}_p(t)$ in Section 3 to inform our choice of the stochastic process for $\bar{p}_p(t)$. Specifically we use two alternative specifications:

- First, we set the value of $\bar{p}_p(t)$ for each portfolio p equal to the unconditional average of the default probabilities shown in Table III.
- Second, we also investigate the role of systematic time-variation in delisting probabilities, by assuming probabilities are instead described by the empirical equation (11).

As we will see however, the low covariance between factors and probabilities renders the practical difference between these two implementations also fairly small.[10]

[9] For the artificial sample it makes no difference whether we report equal or value-weighted returns.

[10] To maximize the possibility of finding any significant differences we allow the probabilities $\bar{p}_p(t)$ to change each month as the risk factors evolve over time.

For each of these cases we then estimate the following reduced form Fama-French 3 factor model to these artificial portfolios[11]

$$r_p(t) = \alpha_p + \beta_{p,m} \times MKT_t + \beta_{p,b} \times HML_t + \beta_{p,s} \times SMB_t + \epsilon_p(t) \qquad (14)$$

and examine the accuracy of our estimated factor loadings and alphas.

4.2 Findings

Table VI shows our findings for the case when the delisting probabilities are constant. Panel A reports the true values of the parameters α_p and β_p used in the data generating process (12). Panel B shows the estimates from the linear factor model (14).

Clearly the most striking result is the finding of large negative alpha's for the last 2 or 3 portfolios where the delisting probabilities are also quite high. These are quantitatively large and, as we show in the next section, very similar to the estimates found in the data.

On the other hand the biases in the factor loadings seem negligible. This is perhaps as expected. We know from equation (6) that when delisting probabilities are constant ($\eta_{f,i} = 0$ there will be dampening effect equal to $1 - \eta_0$ on the estimated coefficients. Since average delisting probabilities (η_0 in this case) are less than 10% for all but the last portfolio this effect is never large.

Table VII documents the changes when the average probability of default covaries (plausibly) with the risk factors. We continue to see a clear pattern of sizable negative alphas on high delistings portfolios. The factor loadings however remain fairly unchanged and only for the last portfolio do we see some signs that the loadings might not be very accurately estimated. It seems then that in practice this second source of bias in estimated factor loadings is not very important either.

Taken together these two tables confirm our impression that linear factor models are likely to lead to potentially important biases when the underlying process is highly non-linear, as is likely the case when we focus on portfolios with many delisting stocks. Most significantly, estimated excess returns can easily, and spuriously, appear large

[11]We omit momentum because it does not alter the conclusions and allows us to reduce the number of reported tables.

and negative for the high delistings portfolios. Biases in factor loadings, however, seem largely irrelevant.

4.3 Empirical Implications

We now discuss possible corrections to allow for proper identification of the underlying parameters of the true stochastic process for expected returns.

We start with the observation that although equations (1) or (4) cannot be directly estimated using linear models on the excess returns $r_i(t)$, we can easily re-write it as:

$$\tilde{r}_i(t) = \frac{r_i(t) - \delta p_i(t)}{1 - p_i(t)} \tag{15}$$

where by definition (equation (1)) the adjusted excess returns $\tilde{r}_i(t)$ now follows the linear factor model[12]

$$\tilde{r}_i(t) = \beta_p F(t) + \epsilon_i(t) \tag{16}$$

Only if the delisting probability equals zero will we have the traditional linear factor regression for unadjusted returns $r_i(t)$, which obtains as a special case of this.

This then suggests adopting the following approach to estimate a factor model on expected equity returns:

1. Estimate the default probability process $\hat{p}_i(t)$ for each firm i

2. Use the estimated probabilities to compute the adjusted excess returns as:

$$\tilde{r}_i(t) = \frac{r_i(t) - \delta \hat{p}_i(t)}{1 - \hat{p}_i(t)} \tag{17}$$

3. Estimate a linear regression of the adjusted excess returns $\tilde{r}_i(t)$ on the pricing factors, $F(t)$.

In practice of course, we will first group stocks into $p = 1, 2, ..P$ portfolios before estimating the factor model, to reduce estimation and measurement errors.

Table VIII shows how this adjustment works in our artificial panel. It reports the results of estimating the same Fama-French 3 factor model (14) but using instead the

[12]Recall that we are normalizing $\alpha_p = 0$ for simplicity

adjusted excess returns in (17). We see that with this adjustment the estimated alphas are now essentially equal to its true value of zero across both sets of portfolios.[13]

Next we investigate the results of implementing this procedure to actual return data.

5 Portfolio Excess Returns

We now use our theoretical insights to re-examine the empirical evidence on distressed stocks. As background, we first report the results of estimating a standard linear factor model on the nine empirical distress portfolios constructed in Section 3. This confirms that our portfolios exhibit the usual pattern of large negative excess returns for distressed portfolios. We then report the results of using the theoretical return correction proposed in the previous section.

5.1 Standard Linear Regressions

Table III documents the basic patterns of stock returns and characteristics across the nine empirical delisting portfolios. We now use these data to estimate monthly excess return regressions for four different empirical models: on an intercept (average excess return), the CAPM, the three-factor Fama-French (1992) regression, and a four-factor Carhart (1997) regression. Recall that, as is common in this literature, we mean excess returns over the return over the market portfolio - as measured by the CRSP VWRETD variable.

The estimated average excess returns for each of these regressions are documented in Panel A of Table IX. They are very much in line with much of the available evidence from other authors. In particular we find that portfolios with high probabilities of delistings - recall that these were fairly negligible for the first 5 or 6 portfolios - average negative excess returns over the market portfolio. Although these are not statistically significant they become much larger once we control for the market and, especially, the Fama-French factors. The last column also shows that much of this "distress" puzzle seems to be linked to the momentum factor. Controlling for this factor reduces

[13] Statistically none is significant at 10%.

abnormal excess returns significantly, to the point where only two portfolios exhibit statistically significant alphas.

For completeness, Panel A in Table X shows the estimated loadings of each portfolio on the market, size and book-to-market factors. As we can see there is a significant size effect across delisting portfolios. Unfortunately however, distressed stocks, which load strongly on size have sizably negative excess returns. Although this pattern of loading more strongly on the risk factors can also be observed for the market and HML the effects are fairly insignificant.

5.2 Non-Linear Model

Panels B for Tables IX and X report the results of incorporating our proposed adjustment that explicitly accounts for the non-linearity introduced by an endogenous probability of delisting.

Formally these panels are constructed from estimating a second set of monthly excess return regressions but where we now adjust the excess return on portfolio p as follows:

$$\tilde{r}_p(t) = \frac{r_p(t) - \bar{p}_p(t)\delta(t)}{1 - \bar{p}_p(t)} \qquad p = 1, 2..9 \tag{18}$$

Each portfolio's probabilities, $\bar{p}_p(t)$, are the equal-weighted averages of the estimated annual delisting firm level probabilities using the logit regression (8) and converted to monthly probabilities using the relation (10). As discussed earlier under the null hypothesis that returns follow the true stochastic process (1) this adjustment correctly removes the non-linear component of returns from the factor regressions, and is suitable to be fitted by a linear factor model.[14]

Panel B in Table IX shows the estimated alphas corrected for delisting bias. These are significantly smaller (in absolute value) than those in Panel A for virtually all models. They are also essentially zero except for a few portfolios when using the Fama-French model. However these middle portfolios are not where we see a significant

[14]The delisting return $\delta(t)$ is time-varying to capture the fact that these returns are often larger in magnitude during recessions. The average correlation between yearly delisting return and yearly predicted delisting probability across portfolios is -0.13.

incidence of default, suggesting these remaining alphas could be driven by something other than financial distress.

Panel B in Table X shows the factor loading estimates from the corrected model. Again, although potentially important in theory these corrections are, in practice, negligible. As we can see the estimates are nearly identical across both panels and continue to exhibit a pronounced size effect across delisting portfolios.

To conclude our results confirm our view that estimation bias is an important driver of the perception that distressed stocks underperform. This result largely, although perhaps not entirely, survive even after we adjust excess returns for various risk factors. Based on this evidence the case for a distress "puzzle" seems considerably weaker. Existing estimates of factor loadings however seem fairly accurate and in particular the conclusion that highly distress stocks load heavily on size is confirmed.

6 Robustness: Distance to Default

6.1 Estimating Delisting Events

Much about our findings hinges on how closely our portfolios capture the available information about delistings related to financial distress. Section 3 shows how closely our estimated probabilities, $\hat{p}_i(t)$, capture actual delistings, at least at the portfolio level where most of the inference about returns is made.

Nevertheless it is useful to examine whether our results are robust to the use of alternative measures of distress related stock delistings. In this section we address this issue by using an alternative estimate of the delisting probabilities $\hat{p}_i(t)$ that do not rely on the reduced form logistic regressions in equation (8). Instead we use the approach suggested by Merton (1974) and estimate a measure of the Moody's Analytics expected distance to default (EDF). The computation details are provided in Appendix C.

Figure 4 shows our average cross-sectional distance to default (DD) estimates and the observed delisting frequencies over the period between 1980 and 2011. The figure shows that our EDF model is able to mimic the movements in delisting probabilities quite closely. There are however two reasons why we did not use this method for our benchmark estimate. The first drawback of DD estimates is apparent from Figure 4: these estimates do not tract the *level* of actual delisting frequencies. The second reason is less obvious but perhaps more significant. Although DD estimates rank stocks relatively well, it is not necessarily able to correctly predict default probabilities in the whole sample or across bins.

To address these two issues we work instead with fitted default probabilities implied by our DD estimates. As before these probabilities are estimated using a logit regression similar to but now with DD as our single predictor of a delisting event:

$$p_i(t) = \frac{1}{1 + \exp^{-\gamma_0 - \gamma_D DD_i(t)}}$$

As before we use the estimated probabilities to construct nine portfolios ranked by delisting probabilities. Figure 5 shows the predicted default probabilities across these portfolios.

6.2 Excess Return Regressions

This section needs to be finished.

7 Conclusion

This paper shows how non-linearities in returns induced by delisting events can affect the inference about the behavior of delisting stocks. Because these events are both extreme and introduce a floor on expected stock returns, the correct factor model is also non-linear. As a result the estimated alphas and loadings in standard linear models are biased. We show that although these biases can be significant for excess returns they are generally quite small for factor loadings. Empirically this occurs largely because the covariance between delisting events and the systematic risk factors is quite small. After we correct these biases we see little evidence of underperformance for portfolios of distressed stocks.

References

Bharath, Sreedhar T., and Tyler Shumway, 2008, Forecasting default with the Merton distance to default model, *Review of Financial Studies* 21,1339-1369.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899-2939.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Center for Research in Security Prices, 2001, CRSP delisting returns, (University of Chicago, Chicago, Ill.).

Dichev, Ilia D., 1998, Is the risk of bankruptcy a systematic risk?, *Journal of Finance*, 53, 1131-1147.

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *Journal of Finance* 55, 389-406.

Fama, Eugene F., and Kenneth R. French,1992, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Garlappi, Lorenzo, Tao Shu, and Hong Yan, 2008, Default risk, shareholder advantage, and stock returns, *Review of Financial Studies* 21, 2743-2778.

Garlappi, Lorenzo, and Hong Yan, 2011, Financial distress and the cross-section of equity returns, *Journal of Finance* 66, 789-822.

Gomes, João F., and Lukas Schmid, 2010, Levered returns, *Journal of Finance* 65, 467-494.

Griffin, John M., and Michael L. Lemmon, 2002, Book-to-Market equity, distress risk, and stock returns, *Journal of Finance* 57, 2317-2336.

Hennessy, Christopher, and Toni Whited, 2007, How costly is external financing? Evidence from a structural estimation, *Journal of Finance* 62, 1705-1745.

Kaplan, Steve, and Jeremy Stein, 1990, How risky is the debt in highly leveraged transactions?, *Journal of Financial Economics*, 27, 215-245.

Korteweg, Arthur, 2004, Financial leverage and expected stock returns: Evidence from pure exchange offers, Working paper, Stanford University.

Leland, Hayne, 1994, Corporate debt value, bond covenants, and optimal capital structure, *Journal of Finance* 49, 1213-1252.

Livdan, Dmitry, Horacio Sapriza, and Lu Zhang, 2009, Financially constrained stock returns, *Journal of Finance* 64, 1827-1862.

Merton, Robert C.,1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449-470.

Shumway, Tyler, 1997, The delisting bias in CRSP data, *Journal of Finance* 52, 327-340.

Shumway, Tyler, 2001, Forecasting Bankruptcy More Accurately: A Simple Hazard Model, *Journal of Business* 74, 101-124.

Shumway, Tyler, and Vincent A. Warther, 2002, The delisting bias in CRSP's Nasdaq data and its implications for the size effect, *Journal of Finance* 54, 2361-2379.

Vassalou, Maria, and Yuhang Xing, 2004, Default risk in equity returns, *Journal of Finance* 59, 831-868.

A Appendix: Delistings

Figure 1 shows the evolution of average delisting returns (both equal and value-weighted) from 1970 to 2011, for the following non-performance delisting codes:

- 500 - Issue stopped trading on exchange - reason unavailable
- 550 - Delisted by current exchange - insufficient number of market makers
- 552 - Delisted by current exchange - price fell below acceptable level
- 560 - Delisted by current exchange - insufficient capital, surplus, and/or equity
- 561 - Delisted by current exchange - insufficient (or non-compliance with rules of) float or assets
- 574 - Delisted by current exchange - bankruptcy, declared insolvent
- 580 - Delisted by current exchange - delinquent in filing, non-payment of fees
- 584 - Delisted by current exchange - does not meet exchanges financial guidelines for continued listing

We remove all delisting returns, $\delta(t)$, greater than positive 100%. In addition all delisting returns are winsorized at the 1-99% to remove outliers. Less than 1% of the delisting returns - out of 5,994 delisting observations - are missing across the whole sample period. This is quite low, when compared to Shumway (1997), and Shumway-Warther (1999), which document about 90% missing data for AMEX-NYSE, and almost all data missing for NASDAQ, or CRSP(2001) which documented the availability of about 73% delisting returns in the 500 series. It seems that CRSP coverage of delisting returns is now almost complete.

The equal-weighted and value-weighted average delisting returns are respectively -28%, and -30%, for the whole sample period, and -36.5%, and -42.3% for the period 2000-2011. The average delisting returns for the whole sample are consistent with Shumway (1997), which reports an average delisting return of -29.9%, for the 1962-1993 sample, for AMEX-NYSE stocks. Interestingly the delisting return $\delta(t)$ seems to exhibit significant time-variation and is often larger in magnitude during market downturns.[15]

Other comments about CRSP delisting codes

- Before 1987, all performance-related and stock-exchange-related delistings were coded 5. After 1987, CRSP started a more refined breakdown. The original code 5 delistings were initially given 500, and are considered to be mainly performance-related delistings (there is only a small number of exchange-related delistings).
- The 572 delisting code (liquidation at company request), is now discontinued and is replaced by the 400 delisting series. The average delisting returns on the 400 series is slightly positive, which may suggest that it does not really reflect negative company performance.
- About 15% of delisting returns reported by CRSP are exactly 0.

[15]The average correlation between yearly delisting return and our yearly predicted delisting probability across portfolios is -0.13.

B Appendix: Firm Level Data and Variables

This appendix describes in detail how our the variables used in the analysis are constructed. All variables codes are for the COMPUSTAT annual file. Quarterly variables typically have a q appended to the end of their variable names.

- Relative size

$$RSIZE_{it} = \log(SIZE_{it}/TOTVAL_t \times 1000)$$

where $TOTVAL_t$ is total dollar value of the S&P500 and

$$SIZE_{it} = PRC_{it} \times SHROUT_{it}/1000$$

- Leverage

$$TLMTA_{it} = LT_{it}/(SIZE_{it} + LT_{it})$$

- Relative cash holdings

$$CASHMTA_{it} = CHE_{it}/(SIZE_{it} + LT_{it})$$

- Market to book ratio

$$MB_{it} = SIZE_{it}/ADJBE_{it}$$

- Adjusted book equity (observation removed if negative)

$$ADJBE_{it} = BE_{it} + 0.1 * (SIZE_{it} - BE_{it})$$

- Stock price

$$PRICE_{it} = \log(\min\{PRC_{it}, 15\})$$

- Excess returns

$$EXRETAVG_{it} = (1-\psi)/(1-\psi^{12}) \times (EXRET_{it} + .. + \psi^{11} EXRET_{it-11})$$

where

$$EXRET_{it} = \log(1 + R_{it}) - \log(1 + VWRETD_t)$$

and $VWRETD_t$ is the return on the S&P500 index. Because of the need for an uninterrupted series any missing variables are set equal to their cross-sectional means.

- Return on assets, or profitability

  $$NIMTAAVG_{it} = (1-\psi^3)/(1-\psi^{12})*(NIMTA_{it,t-2}+\psi^3 NIMTA_{it-3,t-5}+..+\psi^9 NIMTA_{it-9,t-11})$$

  where

  $$\begin{aligned} NIMTAit &= NI_{it}/(SIZE_{it} + LT_{it}) \\ NIMTA_{it-x,t-x-2} &= (NIMTA_{it-x} + NIMTA_{it-x-1} + NIMTA_{it-x-2})/3 \end{aligned}$$

  Because of the need for an uninterrupted series any missing variables are set equal to their cross-sectional means.

- Return volatility

  $$SIGMA_{it} = \sqrt{\frac{252}{N-1}\sum R_{it}^2}$$

  where the summation is of daily returns over the past three months and missing SIGMA observations (when $N < 5$) are replaced with the cross-sectional mean.

and we use $\psi = 2^{-1/3}$. Each of these variables is also winsorized at the fifth and ninety-fifth percentiles across all firm-months. Furthermore, following Campbell et al. (2008), all observations with missing size, profitability, leverage, or excess return data are dropped.

C Appendix: Distance to Default

C.1 Theory

As Merton (1974) shows, the expected distance-to-default for a firm with liabilities D with maturity T and enterprise value V is given by the formula:

$$DD = \frac{\log \frac{V}{D} + (\mu - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \tag{19}$$

This assumes that default occurs if $V < D$ at time T. In (19), μ and σ are the drift and instantaneous standard deviation in the stochastic process for asset value, which follows a geometric Brownian motion.

C.2 Data and Implementation

We consider the sample period 1980 to 2011, which includes 227,237 firm-year observations, and 4801 delistings[16] The EDF estimates are derived annually based on December's estimations.

In order to compute the probability of default as defined above, we use an implied equity premium of 6% so that the drift term is $\mu = r + 0.06$ and solve for the implied value of the assets of the firm V and its volatility σ using the follow iterative procedure:

- Define $E = SIZE$ as the observed market value of firm equity and let $\sigma_e = SIGMA$ be the standard deviation of equity returns.
- As it is common in the literature define the market value of debt as:

$$D = \underbrace{DLC}_{\text{Current debt}} + 0.5 \underbrace{DLTT}_{\text{LT debt}}$$

  and let $V = V_0 = E + D$

- For each year t, solve for the fixed point problem associated with the volatility of total asset returns, using daily data spanning the period [t-1, t):
  1. Start with initial guess $V = V_0$, and construct asset volatility $\sigma = \sigma_0 = \frac{V_0}{E}\sigma_e$.
  2. Use these estimated values for σ_0 and the data on r, D, and E in the Black-Scholes formula to obtain the implied valued of the firm V_1

[16]For 1970-2011, we have 268,546, and 5,332 observations, respectively.

3. Given the updated firm value V_1, revise estimated asset volatility $\sigma = \sigma_1$
4. Iterate until convergence of σ and V.

Data Adjustments:

We apply the following data adjustments in order to solve for the fixed point problem for the largest possible sample:

- When data on D is missing, we use $D = \text{median}(D/(LTQ)) * (LTQ)$ where the median is calculated across observations with small non-zeros values of D. We also replace missing LTQ data by median(LTQ).
- When D or E are large ($> 100,000$), we use $D = D/10000$, and $E = E/10000$.
- We also floor the level of σ at 5%.

Table I: **Summary Statistics**

This table reports summary statistics for the core variables used in the logistic regressions. The data are monthly over the period 1950 to 2011.

Variable	N	Mean	Median	Std Dev	Minimum	Maximum
NIMTA	2,611,567	0.000	0.006	0.034	-0.201	0.056
TLMTA	2,611,545	0.431	0.405	0.278	0.008	0.966
EXRET	2,598,310	-0.009	-0.008	0.139	-0.471	0.432
RSIZE	2,611,576	-10.277	-10.396	2.071	-14.661	-5.156
SIGMA	2,611,580	0.542	0.434	0.381	0.099	2.180
CASHMTA	2,592,000	0.094	0.047	0.128	0.000	0.748
MB	2,611,580	2.015	1.480	1.724	0.204	9.821
PRICE	2,611,580	2.031	2.555	0.961	-1.505	2.708

Table II: **Logistic Regression Estimates**

This table reports the estimated coefficients for the full sample period (data up to December 2011) for the logistic regression (8).

Parameter	**Estimate**	**Std Err**	**p-value**
CONSTANT	-9.794	0.235	<0.0001
EXRETAVG	-6.177	0.270	<0.0001
SIGMA	0.326	0.035	<0.0001
MB	0.174	0.010	<0.0001
NIMTAAVG	-7.831	0.419	<0.0001
TLMTA	0.978	0.071	<0.0001
CASHMTA	-0.909	0.111	<0.0001
RSIZE	-0.441	0.018	<0.0001
PRICE	-0.579	0.021	<0.0001
N	199,904	**Pseudo R-squared**	0.389
N - Delisting	4,440		

Table III: **Properties of Delisting Portfolios**

This table reports summary statistics for the portfolios constructed using the estimated probabilities of default using the logistic regression (8). Excess returns are over the market portfolio, defined as the S&P500 index. This data covers monthly data from 1970 until 2011. Some denoted quantities are annualized.

PORTFOLIO	Annual Pr(default)	RSIZE	MB	Annual average excess return	Skewness	Annual standard deviation
1	0.0003	-5.61	2.41	-0.001	0.438	0.026
2	0.0006	-6.89	2.25	0.008	0.757	0.060
3	0.0009	-7.65	2.23	0.018	0.062	0.071
4	0.0020	-8.53	2.35	0.017	0.092	0.100
5	0.0049	-9.49	2.46	0.001	0.646	0.129
6	0.0151	-10.37	2.60	-0.001	1.383	0.178
7	0.0431	-11.11	2.93	-0.027	1.523	0.234
8	0.0844	-11.57	3.26	-0.031	2.010	0.300
9	0.1669	-11.90	3.38	-0.044	2.334	0.364

Table IV: **Actual and Estimated Delisting Frequencies**

This table reports R^2 associated with regressing ex-post default frequencies on the average estimated probabilities of default for nine portfolios. Each portfolio is constructed using the estimated default probabilities using the logistic regression (8). Logit regressions coefficients are calculated in December and applied over the entire following year. The predicted probability estimate over the entire following calendar year is paired with the year's corresponding realized default frequency. This annual data is from 1970 until 2011.

PORTFOLIO	**R-squared**	**Coefficient**
1	0.018	1.802
2	0.039	0.258
3	0.071	0.393
4	0.533	0.958
5	0.662	0.962
6	0.854	1.031
7	0.849	1.083
8	0.862	1.153
9	0.901	1.039

Table V: **Systematic Components of Delisting Probabilities**

This table reports the results of regressing the average estimated probabilities of default for nine portfolios on high order polynomials of the Fama-French risk factors. Each portfolio is constructed using the estimated default probabilities using the logistic regression (8). Sample period is from 1970 to 2011 at a monthly frequency.

	PORTFOLIO								
Loadings	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
CONS	0.00**	0.00**	0.00**	0.00**	0.00**	0.01**	0.04**	0.08**	0.16**
MKT	0.00	0.00	0.00	0.00	0.00	0.01	0.05	0.09	0.22
SMB	0.00*	0.00	0.00	0.00	0.00	-0.01	-0.05	-0.09	-0.14
HML	0.00	0.00	0.00	0.00	-0.02*	-0.06*	-0.17*	-0.21	-0.19
MOM	0.00	0.00	0.00	0.00	-0.01	-0.02	-0.06	-0.13	-0.29*
SMB*HML	-0.01	-0.01	-0.01	-0.07	-0.24	-0.54	-1.12	0.31	2.49
SMB*MOM	0.00	0.00	0.00	0.00	-0.09	-0.41	-1.44	-1.60	-3.11
HML*MOM	0.00	0.01	0.01	0.02	0.02	0.13	1.27	4.29**	9.16**
MKT2	-0.01	0.00	-0.01	-0.02	-0.03	-0.03	-0.12	-0.40	-1.81
SMB2	-0.01	0.00	-0.01	-0.05	-0.13	-0.36	-0.81	-1.11	-1.20
HML2	0.02	0.02*	0.06**	0.17**	0.32*	0.53	1.52	3.39	7.73*
MOM2	0.01*	0.02**	0.03**	0.08**	0.21**	0.60**	1.65**	3.11**	4.69**
MKT3	-0.05	-0.02	-0.10	-0.20	-0.37	-1.50	-5.30	-9.55	-20.48
SMB3	0.00	-0.04	-0.09	-0.25	-0.61	-0.70	-0.49	1.45	8.42
HML3	-0.05	0.05	0.16	0.50	1.26	3.82	2.42	-12.84	-40.40
MOM3	0.04*	0.04*	0.08*	0.19*	0.55**	1.35*	3.62	5.13	7.90
R-squared	0.05	0.10	0.10	0.13	0.12	0.08	0.09	0.09	0.10

** - 1% significance, * - 5% significance

Table VI: **Simulation Results: Constant Delisting Probabilities**

This table reports the results of estimating a linear Fama-French 3 factor model on our simulated portfolios when the delisting probabilities are constant and equal to their time series averages for each portfolio. Panel A shows the true factor loadings in the data generating process (1). Panel B reports the empirical estimates from (14).

PANEL A

True Model Parameters				
PORTFOLIO	**ALPHA**	**MKT**	**SMB**	**HML**
1	0	-0.054	-0.067	-0.067
2	0	0.025	0.325	-0.024
3	0	0.063	0.465	0.076
4	0	0.073	0.746	0.037
5	0	0.082	1.042	0.134
6	0	0.172	1.333	0.220
7	0	0.213	1.595	0.155
8	0	0.204	1.897	0.096
9	0	0.130	2.213	0.103

PANEL B

Standard FF Regression Estimates				
PORTFOLIO	**ALPHA**	**MKT**	**SMB**	**HML**
1	-0.005	-0.048	-0.045	-0.068
2	0.000	0.023	0.318	-0.042
3	0.002	0.057	0.471	0.075
4	-0.009	0.083	0.729	0.043
5	-0.009	0.084	1.020	0.136
6	-0.020	0.171	1.320	0.188
7	-0.044	0.214	1.573	0.148
8	-0.087	0.201	1.877	0.092
9	-0.191	0.148	2.170	0.101

Table VII: **Simulation Results: Time Varying Delisting Probabilities**

This table reports the results of estimating a linear Fama-French 3 factor model on our simulated portfolios when the delisting probabilities covary with the risk factors according to the estimates from equation (11). Panel A shows the true factor loadings in the data generating process (1). Panel B reports the empirical estimates from (14).

PANEL A

True Model Parameters				
PORTFOLIO	**ALPHA**	**MKT**	**SMB**	**HML**
1	0	-0.054	-0.067	-0.067
2	0	0.025	0.325	-0.024
3	0	0.063	0.465	0.076
4	0	0.073	0.746	0.037
5	0	0.082	1.042	0.134
6	0	0.172	1.333	0.220
7	0	0.213	1.595	0.155
8	0	0.204	1.897	0.096
9	0	0.130	2.213	0.103

PANEL B

Standard FF regression				
PORTFOLIO	**ALPHA**	**MKT**	**SMB**	**HML**
1	0.009	-0.060	-0.061	-0.069
2	-0.006	0.001	0.322	-0.064
3	0.009	0.081	0.451	0.086
4	-0.006	0.063	0.746	0.033
5	-0.007	0.080	1.042	0.146
6	-0.008	0.161	1.315	0.216
7	-0.021	0.196	1.559	0.157
8	-0.042	0.180	1.886	0.098
9	-0.073	0.086	2.133	0.033

Table VIII: **Simulation Results: Estimating Adjusted Returns**

This table reports the results of estimating a linear Fama-French 3 factor model on our simulated portfolios to the excess returns, adjusted according to (17). Panel A shows the results when the underlying true delisting probabilities are constant. Panel B shows the results when the underlying true delisting probabilities covary with the risk factors.

PANEL A

Constant Probabilities				
PORTFOLIO	**ALPHA**	**MKT**	**SMB**	**HML**
1	-0.004	-0.048	-0.045	-0.068
2	0.000	0.023	0.318	-0.042
3	0.003	0.057	0.471	0.075
4	-0.007	0.083	0.729	0.043
5	-0.004	0.084	1.021	0.136
6	-0.005	0.172	1.322	0.189
7	0.000	0.214	1.579	0.149
8	0.001	0.202	1.891	0.092
9	-0.010	0.151	2.203	0.103

PANEL B

Time Varying Probabilities				
PORTFOLIO	**ALPHA**	**MKT**	**SMB**	**HML**
1	0.009	-0.060	-0.061	-0.069
2	-0.006	0.001	0.322	-0.064
3	0.010	0.081	0.451	0.087
4	-0.005	0.063	0.747	0.033
5	-0.005	0.081	1.043	0.147
6	-0.002	0.163	1.318	0.217
7	-0.001	0.202	1.566	0.158
8	-0.003	0.195	1.908	0.107
9	0.005	0.120	2.192	0.073

Table IX: **Excess Returns Across Delisting Portfolios**

This table reports the average excess returns overt the market portfolio as well as the excess return over three different empirical models: the CAPM, the three-factor Fama-French (1992) regression, and a four-factor Carhart (1997) regression. Panel A uses raw portfolio returns while Panel B adjusts portfolio returns for delisting events. Each portfolio is constructed using the estimated default probabilities using the logistic regression (8). Sample period runs monthly from 1970 until 2011.

PANEL A

Standard FF regression				
PORTFOLIO	**CONSTANT**	**CAPM**	**FF 3-FACTOR**	**CARHART 4-FACTOR**
1	-0.001	0.002	0.007*	0.000
2	0.008	0.002	-0.001	0.001
3	0.018	0.010	-0.001	0.005
4	0.017	0.005	-0.008	0.000
5	0.011	-0.004	-0.026***	-0.012
6	-0.005	-0.027	-0.058***	-0.027**
7	-0.027	-0.055	-0.086***	-0.049**
8	-0.031	-0.062	-0.095***	-0.046
9	-0.044	-0.074	-0.111**	-0.064

*** - 1% significance, ** - 5% significance, * - 10% significance

PANEL B

Adjusted return regression				
PORTFOLIO	**CONSTANT**	**CAPM**	**FF 3-FACTOR**	**CARHART 4-FACTOR**
1	-0.001	0.002	0.007*	0.000
2	0.008	0.003	-0.001	0.001
3	0.018	0.010	0.000	0.005
4	0.017	0.005	-0.007	0.001
5	0.012	-0.002	-0.025***	-0.010
6	0.000	-0.022	-0.053***	-0.022
7	-0.013	-0.040	-0.072***	-0.033
8	-0.001	-0.032	-0.064*	-0.013
9	0.024	-0.006	-0.044	0.010

*** - 1% significance, ** - 5% significance, * - 10% significance

Table X: **Factor Loadings Across Delisting Portfolios**

This table reports the loading on the market, size and book-to-market factors for the excess returns overt the market portfolio for each delisting portfolio. Panel A uses raw portfolio returns while Panel B adjusts portfolio returns for delisting events. Each portfolio is constructed using the estimated default probabilities using the logistic regression (8). Sample period runs monthly from 1970 to 2011.

Panel A

Standard FF regression

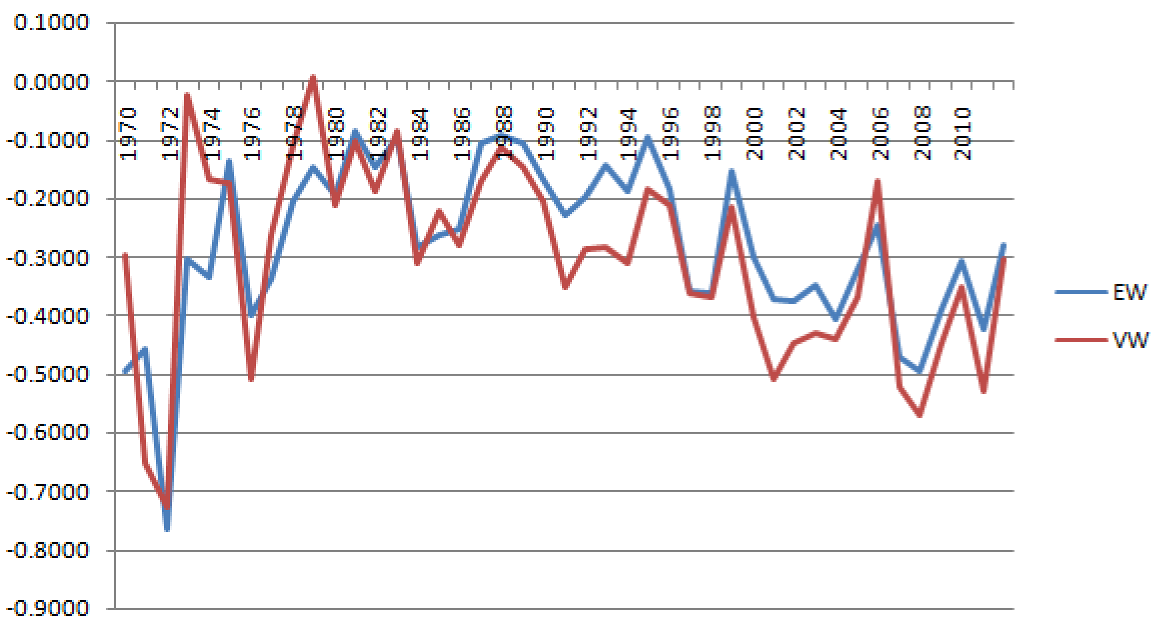
PORTFOLIO	**MKT**	**HML**	**SMB**
1	-0.054	-0.067	-0.063
2	0.025	-0.024	0.325
3	0.063	0.076	0.465
4	0.073	0.037	0.746
5	0.082	0.134	1.042
6	0.171	0.220	1.330
7	0.210	0.158	1.587
8	0.199	0.102	1.877
9	0.126	0.117	2.159

Panel B

Adjusted return regression

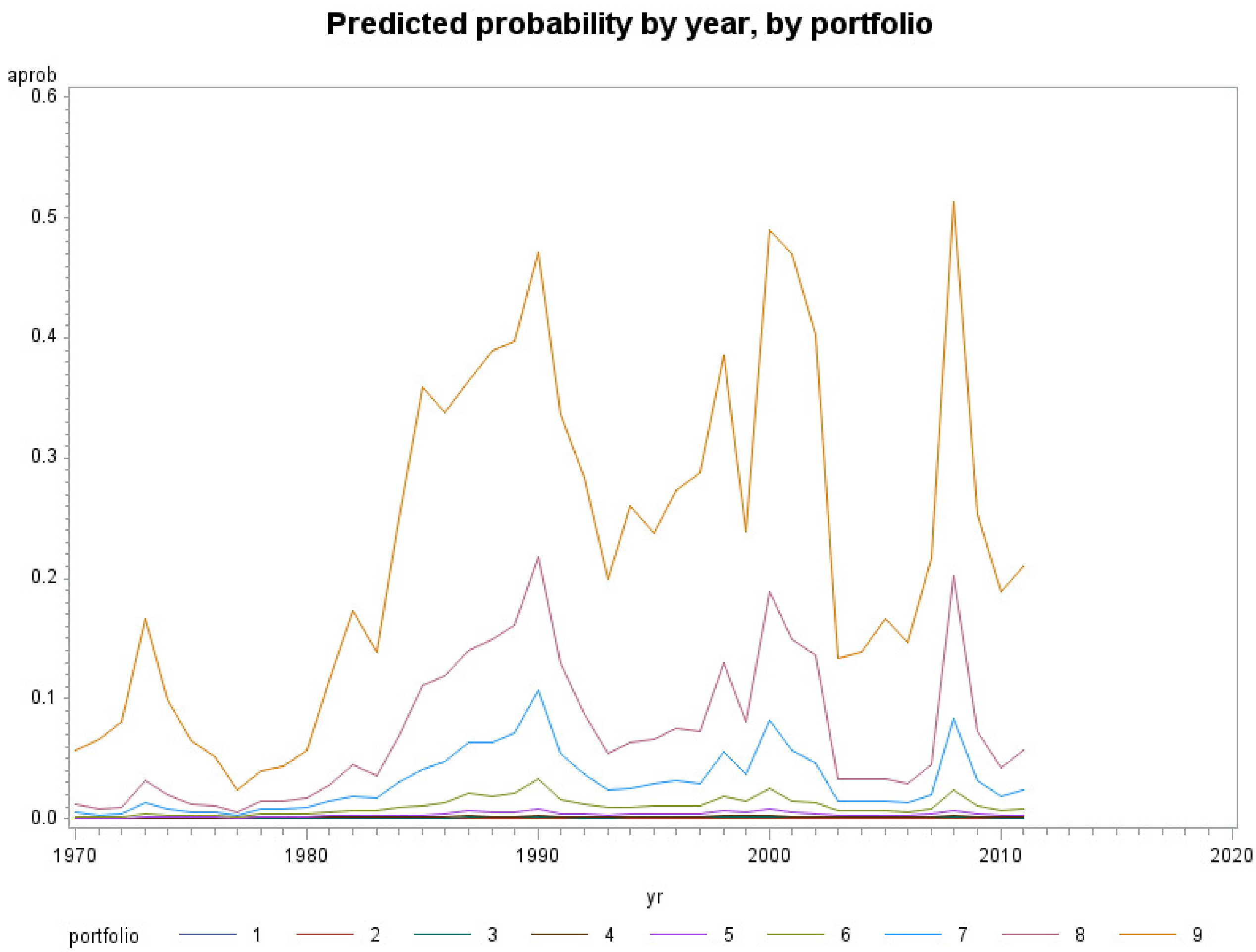
PORTFOLIO	**MKT**	**HML**	**SMB**
1	-0.054	-0.067	-0.063
2	0.025	-0.024	0.325
3	0.063	0.076	0.465
4	0.073	0.037	0.746
5	0.082	0.134	1.042
6	0.172	0.220	1.333
7	0.213	0.155	1.595
8	0.204	0.096	1.897
9	0.130	0.103	2.213

Figure 1: **Delisting Returns**



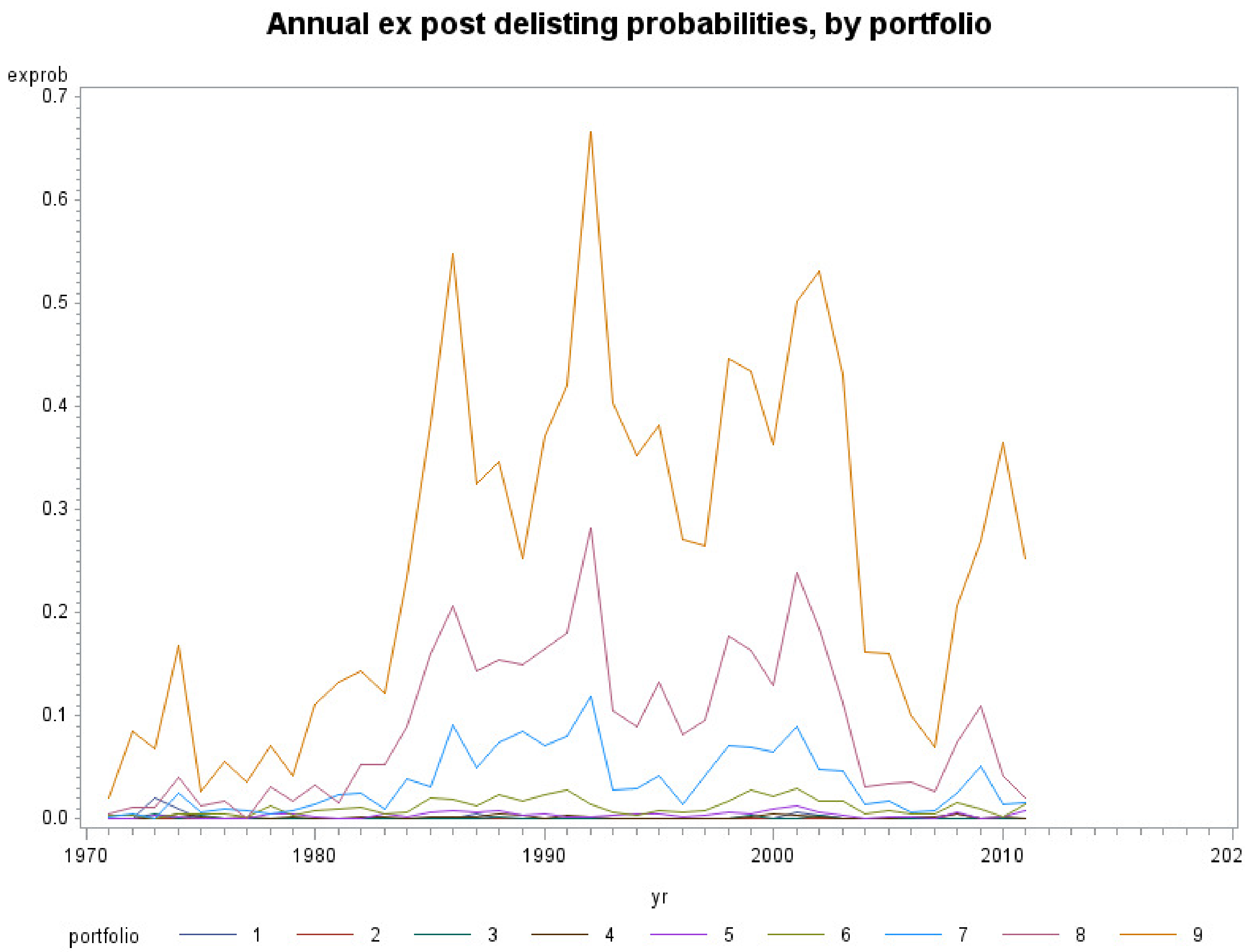
This figure shows the observed delisting returns over the period 1970-2011.

Figure 2: **Estimated Probabilities for the Logistic Model**



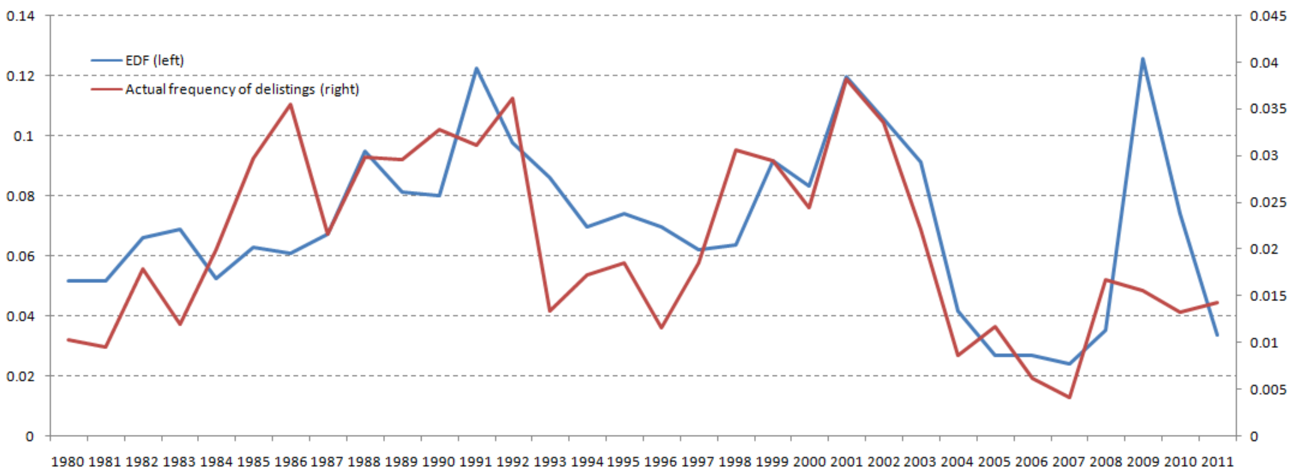
This figure shows the estimated delisting probabilities $\hat{p}_i(t)$ from the benchmark logistic model in equation 8).

Figure 3: **Estimated Probabilities for the Logistic Model**



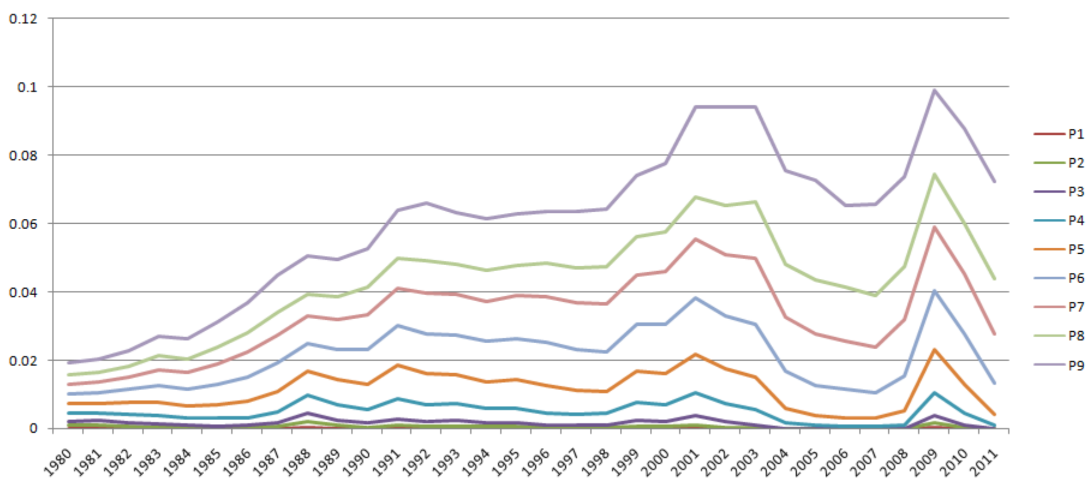
This figure shows the ex-post delisting frequencies for the 9 portfolios constructed using the estimated default probabilities, $\hat{p}_i(t)$, using the benchmark logistic model in equation 8.

Figure 4: **Expected Default Frequencies and Actual Delistings**



This figure shows the average cross-sectional estimated EDF (left axis) and the actual delistings (right axis) over the period 1980-2011.

Figure 5: **Predicted Probabilities of Default, per EDF portfolio, 1980-2011**



This figure shows the average cross-sectional estimated EDF (left axis) and the actual delistings (right axis) over the period 1980-2011.