

Death and Jackpot: Why Do Individual Investors Hold Overpriced Stocks?

Jennifer Conrad*, Nishad Kapadia⁺, and Yuhang Xing⁺

This draft: February 2013

Abstract

Campbell, Hilscher, and Szilagyi (2008) show that firms with a high probability of default have abnormally low average future returns. We show that firms which have a high potential for default ('death') also tend to have a relatively high probability of extremely large ('jackpot') payoffs. We develop a model to predict jackpot returns and find that stocks with high predicted probabilities for jackpot returns earn abnormally low average returns consistent with an investor preference for skewed, lottery-like payoffs. Stocks with high death or jackpot probabilities have relatively low institutional ownership and the jackpot effect we document is much stronger in stocks with high limits to arbitrage. Jackpot and death probabilities are significantly correlated across firms, but we find evidence that jackpot effects are not merely a reflection of default effects: firms with high default probability but low jackpot probability (those with low past return skewness and low market-to-book ratios and sales growth) do not exhibit low average returns and both effects remain significant in Fama-MacBeth regressions. Overall, we find evidence consistent with the hypothesis that individual investors hold 'overpriced' stocks due to a preference for skewness and that the low average returns of such stocks persist due to limits to arbitrage.

JEL: G11, G12, G32, G33

Keywords: Distress risk, skewness, stock returns, anomalies

* Kenan-Flagler Business School, UNC-Chapel Hill, McColl 4109, CB #3490, Chapel Hill, NC 27599.

⁺ Jones Graduate School of Business, Rice University, 6100 Main Street, Houston TX 77005.

We thank an anonymous referee, Nicholas Barberis, Sebastien Michenaud, and seminar participants at the China International Conference in Finance and Rice University for helpful comments. Contact author: Nishad Kapadia. Email: nishadk@rice.edu

Death and Jackpot: Why Do Individual Investors hold Overpriced Stocks?

Abstract

Campbell, Hilscher, and Szilagyi (2008) show that firms with a high probability of default have abnormally low average future returns. We show that firms which have a high potential for default ('death') also tend to have a relatively high probability of extremely large ('jackpot') payoffs. We develop a model to predict jackpot returns and find that stocks with high predicted probabilities for jackpot returns earn abnormally low average returns consistent with an investor preference for skewed, lottery-like payoffs. Stocks with high death or jackpot probabilities have relatively low institutional ownership and the jackpot effect we document is much stronger in stocks with high limits to arbitrage. Jackpot and death probabilities are significantly correlated across firms, but we find evidence that jackpot effects are not merely a reflection of default effects: firms with high default probability but low jackpot probability (those with low past return skewness and low market-to-book ratios and sales growth) do not exhibit low average returns and both effects remain significant in Fama-MacBeth regressions. Overall, we find evidence consistent with the hypothesis that individual investors hold 'overpriced' stocks due to a preference for skewness; we also find evidence that the low average returns of such stocks persist due to limits to arbitrage.

1 Introduction

Campbell, Hilscher and Szilagyi (2008; henceforth CHS) present convincing evidence that stocks with a high probability of default have low subsequent returns. This result is puzzling; as they point out, it suggests that the market has not priced distress risk appropriately. In considering possible explanations for this anomaly, CHS note that individual securities with high failure probabilities, as well as portfolios of distressed stocks, have positive skewness; they conjecture that investors with a strong preference for positive skewness may bid up the prices of these securities, leading to low subsequent returns.

This conjecture is a natural one given recent papers on the relation between skewness and returns. Barberis and Huang (2008; henceforth BH) consider an economy in which investors have cumulative prospect theory preferences. They show that, in such an economy, positively skewed securities can become overpriced and earn negative average excess returns. There is also growing empirical evidence that securities with positive skewness or a high probability of extreme positive outcomes ('jackpots')¹ have low subsequent returns (see, e.g., Mitton and Vorkink (2007), Kumar (2009), Boyer, Mitton and Vorkink (2010), Bali, Cakici and Whitelaw (2011) and Conrad, Dittmar and Ghysels (2012)). However, both CHS and BH argue that for such an effect to persist, there must be risks or costs that prevent other rational investors (with potentially different preferences) from arbitraging this effect away. Thus there are two important questions regarding the low average returns of distressed stocks. First, is it skewness that motivates investors to hold these stocks despite their low expected returns, and second, why are these low returns not arbitrated away? In this paper, we focus primarily on understanding the

¹ Throughout the paper, we call these extreme positive outcomes 'jackpots' or jackpot returns, following Barberis and Huang (2008).

first question, by examining whether skewed, lottery-like payoffs make stocks with high default probabilities attractive to a segment of investors, resulting in low average returns; similar to CHS, we provide some evidence on the second question by showing that the jackpot effect we find is strongest in stocks for which arbitrage trades may be relatively costly.

The connection between high default risk and high jackpot probability can be understood in the context of the Merton (1974) model, which views equity as a call option on the assets of the firm. For firms close to the default boundary, this optionality is more important, leading to more skewed pay-off distributions. We begin by confirming that distressed stocks have lottery-like payoffs. We go on to perform a calibration, based on the BH model, and show that the distribution of returns of highly distressed stocks is sufficiently lottery-like that a heterogeneous-holdings equilibrium exists in which these stocks have negative expected returns. The predicted returns for high default probability stocks from the calibrated BH model match the low returns we observe in the data.

To investigate the relation between the effects of high default probabilities and the probability of extreme positive returns, we develop measures of both for individual securities. Specifically, following CHS, we use a logit model to predict both distress ('death') and extreme positive payoffs ('jackpots'). This framework is also similar in spirit to that of BH, who model skewed securities as having binary payoffs, with jackpot payoffs arising with some probability q . In our model, we use several benchmarks for the payoff to a security to be considered as a jackpot. We find that stocks with high predicted probability of having a jackpot return subsequently earn low average returns and have negative four-factor alphas, with magnitudes similar to stocks with high default probability. We also find that *ex-ante* probabilities of death and jackpots are highly correlated, with pair-wise correlations of approximately 40% in our

sample. More than 50% of the firms in the highest quintile of predicted distress are also in the highest quintile of predicted jackpot. This correlation does not arise from volatility, since we find that the probability of death is a far better predictor of future jackpots than using volatility alone. Thus the jackpot and the death effects are closely linked.

Since prior research argues that the preference for skewness is more likely to be displayed by retail investors than institutions (Kumar (2009), BH), we examine the ownership structure for stocks with high default and high jackpot probability. We find that the degree of institutional ownership declines significantly as we move to higher default and jackpot probability deciles in the cross-section. Over time, institutional ownership has increased in US markets, particularly in smaller stocks as institutions search for “greener pastures” (Bennet, Sias, and Starks (2003)). However, we find that compared to the median stock, institutional ownership has increased at a slower rate for stocks in both the highest death and highest jackpot deciles. Finally, using an event-study, we show that the level of institutional ownership in a firm declines from four quarters prior to the firm’s entry into the top death or jackpot probability decile and that this decline continues for the four quarters subsequent to entry into the decile. Thus, the ownership structure of both high death and high jackpot probability stocks is consistent with the hypothesis that these stocks are largely owned by retail investors, who are more likely to display a preference for skewness.

We also examine the sensitivity of jackpot stocks’ low returns to three measures of limits to arbitrage: size, residual analyst coverage, and residual institutional ownership (where both residuals are calculated by controlling for size). We find that the abnormally low returns of jackpot stocks are statistically and economically significant among securities where limits to arbitrage are expected to be high: in small market-capitalization, low residual institutional

ownership and (weakly) low analyst coverage firms. Abnormal returns are reduced substantially in magnitude (typically to half or less of their original value) and are statistically insignificant for all measures when limits to arbitrage are low. Thus, low returns to jackpot stocks are associated with high limits to arbitrage in such stocks.

Finally we examine the relation between the effects of jackpot probability and default probability on expected returns. That is, we test whether the jackpot effect is separate from the effect of default probabilities. Jackpot and default probabilities are sufficiently strongly correlated that we begin with an indirect test of whether high default probability firms have low returns when the probability of a jackpot payoff is relatively low.² To proxy for predicted jackpots, we examine two sets of variables: past return skewness and two variables related to growth – market-to-book ratio and sales growth. We find that the low returns documented by CHS are concentrated in those firms in the top 30% of realized daily log return skewness over the past three months and are not present in firms that comprise the bottom 30% of realized past skewness. We also sort firms in the highest predicted distress quintile into ‘traditional distressed firms’ and ‘speculative distressed firms’. Traditional distressed firms are defined as firms with low sales growth and low market-to-book ratios (in the bottom 30% of the full sample for both) and speculative distressed firms are those with high market-to-book ratios and high sales growth (in the top 30% of the full sample for both). The average probability of distress is high (in the top quintile by construction) for both these sets of firms, however speculative distress firms have almost twice the probability of a jackpot payoff as the traditional distress firms. We find that

² Double-sorting on the basis of default and jackpot probabilities generated portfolios which had too few securities to draw reliable inferences. We also examined sub-samples of securities with zero debt, and high jackpot probabilities.

only the speculative firms have low subsequent average returns, with four factor alphas of -1.7% a month, while the alpha for the traditional distress firms is not statistically different from zero. This result suggests that the effect of a high jackpot probability—or positive skewness-- on subsequent returns is not merely a different way to measure the effect of a high distress probability on returns.

We also report results from Fama-MacBeth regressions that allow us to test whether both default and jackpot probabilities are important in predicting returns while controlling simultaneously for other variables that have been shown to affect expected return and are correlated with both these measures. Of course, given the relatively strong relation between default and jackpots, we cannot rule out the possibility that measurement error in jackpot and distress probabilities may influence our results. However, we find that both distress and jackpots have a significant impact on expected returns after including standard controls such as past volatility, market capitalization and book-to-market equity ratio; indeed, the effect of a one-standard deviation shock to both jackpot and distress generates an effect of very similar magnitude on returns.. This result also suggests that distress and jackpots may have distinct effects, with both having an impact on pricing and, as a consequence, on subsequent returns.³

Overall, our results are consistent with the hypothesis that a preference for lottery-like payoffs explains why rational investors (individuals with prospect theory based utility functions) hold on to distressed stocks even though they have low average returns, with these low returns persisting due to limits to arbitrage. As discussed above, these results are consistent with the

³ In a simulation exercise, we show that in the presence of measurement error, Fama-MacBeth regressions containing two correlated independent variables will find both variables to be significant, even if only one is truly correlated with expected returns.

literature that finds that assets with lottery-like payoffs earn low average returns. The findings are also consistent with evidence from the option pricing literature. Ni (2009) finds that out-of-the-money call options (with highly skewed return distributions) have significant negative average returns. Ni and Lemmon (2009) show that in sharp contrast to index options, which are mostly traded by institutional investors, individual stock options are mostly traded by individual investors. More recently, Boyer and Vorkink (2011) show that skewness in option markets is priced in a manner consistent with a preference for total skewness; options with greater skewness earn lower average returns. All these findings support the hypothesis that a preference for skewed assets has a significant impact on asset pricing.

Our paper is also related to the extensive literature that examines the relation between distress risk and expected stock returns. Fama and French (1996) and Vassalou and Xing (2004) argue that there is a positive relation between distress risk and expected returns. In contrast, Dichev (1998) finds that firms with high ‘O-scores’, from the Ohlson (1980) model, have low average returns. Griffin and Lemmon (2002) argue that the results in Dichev (1998) are driven by mispricing amongst high distress risk stocks with high M/B ratios. This result is similar to our evidence that the low returns to stocks with high default probability, as documented by CHS, are visible in speculative stocks but not in traditional distress stocks. We also show that speculative stocks have much larger probabilities for a jackpot payoff than traditional distressed stocks, providing an economic rationale for their high valuations.

Recent explanations for the CHS results include Chava and Purnandam (2010), George and Hwang (2010), and Garlappi and Yan (2011), who argue that small-sample effects, the costs of financial distress, and differences in shareholder recovery, respectively, are responsible for the low average returns of high default probability stocks. However, Gao, Parsons, and Shen (2012)

find that the distress effect is present in a sample of 39 countries and is not related to the extent of creditor protection. Another hypothesis is that idiosyncratic volatility is responsible for the low returns of stocks with high default probabilities. In their tests, CHS report that this does not appear to hold, since they find that other variables in their model besides volatility are important in predicting the low returns of such stocks. In our own tests, we find that both default probability and the probability of a jackpot payoff retain their statistical and economic significance in predicting returns in Fama-MacBeth regressions *after* controlling for volatility. In fact, these variables render idiosyncratic volatility insignificant in these regressions.

Our paper is organized as follows: in Section 2 we motivate our paper by showing that stocks with high default probability have lottery-like payoffs and calibrate the Barberis and Huang (2008) model to examine whether such payoffs would earn negative average returns in the model. Section 3 describes the model for estimating the probability of a jackpot payoff; in Section 4, we examine whether a potential for a jackpot return can explain the low average returns of stocks with high default risk; in Section 5 we provide additional results on the relation between the effects of distress and jackpots on expected returns; Section 6 includes further robustness checks and we conclude in Section 7.

2 Motivation: The link between default probability and jackpots

In this section, we motivate our conjecture that the probability of jackpot returns drives the default risk effect, by examining whether sorts on default probability also results in sorting on jackpot returns. We then perform a calibration exercise, based on the BH model, in order to examine whether stocks with high default probability have sufficient skewness for the BH equilibrium to exist.

2.1 Are distressed stocks likely to have jackpots pay-offs?

This section examines the link between default probability and the probability of earning jackpot returns. We first construct a measure of default probability (DEATHP) of each stock from the model in CHS (Table IV, 12 month lag, page 2913 of CHS). This measure is constructed as described in CHS and requires quarterly CompuStat data, restricting the sample from 1972-2009. We recalculate DEATHP every month and sort stocks into deciles based on DEATHP. Table 1, Panel A examines the properties of decile portfolios based on DEATHP. We skip a month between computing DEATHP and measuring returns, to ensure that our results are not driven by short-term reversals. We first show that the findings in CHS extend to our sample: stocks with high predicted default probability have low returns and low Carhart four factor alphas. The difference between the safest and the most likely to default decile is about -1% per month for both returns and four factor alphas. Thus, stocks with high default probability seem singularly unattractive to investors. However, they have one redeeming feature. As the next three rows of Table 1 show, these stocks have high probabilities of delivering jackpot returns. For example, the fraction of stocks in the safest portfolio that has log returns greater than 100% over the next year is 1.4%. This fraction almost triples, to 3.9%, for the portfolio with the highest default probability. Similarly, the average skewness of daily stock returns in the safest portfolio decile over the next year is 0.68. The average skewness observed in the portfolio with the highest default probability increases to 1.58. We also examine the time-series average of median institutional ownership for these portfolios. We can see that stocks with high DEATHP have low institutional ownership. This suggests that these stocks are largely owned by retail investors, who are more likely to display a preference for skewness (Kumar (2009)). These stocks are also likely

to have high arbitrage costs, consistent with the arguments in CHS and Barberis and Huang (2008).

Barberis and Huang (2008) show that an asset with jackpot returns is so attractive to investors with prospect theory utility functions, that a heterogeneous holding equilibrium exists in which such investors are indifferent between holding the market portfolio and an under-diversified portfolio in which the asset with jackpot returns has a non-trivial weight. The asset with jackpot returns earns negative expected returns in this equilibrium. However, for the equilibrium to exist, the pay-offs of the jackpot asset must be sufficiently skewed. We calibrate the Barberis and Huang (2008) model to test whether stocks with high distress risk have sufficient skewness for the heterogeneous holding equilibrium to exist.

2.2 Calibrating the Barberis and Huang (2008) model⁴

In Barberis and Huang (2008), the jackpot asset has a binary payoff structure $(L, q ; l, 1-q)$, earning a gross pay-off of L with probability q , and a payoff of l with probability $1-q$. To closely mimic this setup, we model the pay-off of the typical stock in the highest distress risk decile portfolio as binary. We set a jackpot return as a log return greater than 100% over the next year. Thus, the likelihood of a jackpot return is $q = Probability\ of\ Log(R_{i,t}) > 1$, where $R_{i,t}$ is the gross return of stock i , at time t , in the highest distress risk portfolio.

In the jackpot state, the expected gross return of a typical stock in this portfolio with price ' P ' is L/P . Given our definition of a jackpot return, this return is $E(R_{i,t} | Log(R_{i,t}) > 1)$. Similarly the expected return in the bad state is $l/P = E(R_{i,t} | Log(R_{i,t}) \leq 1)$. We measure q ,

⁴ We would like to thank an anonymous referee for suggesting that we calibrate the Barberis and Huang (2008) model.

L/P and l/P using their sample means from the pooled sample of all stocks in the highest distress decile portfolio from 1972 to 2009. We find that $\hat{q} = 3.9\%$, $\widehat{L/P}=4.24$, and $\widehat{l/P}=0.91$.

In order to see if this pay-off structure will support the heterogeneous-holdings equilibrium, we solve the BH model for different values of q , using their assumptions regarding values for parameters of the utility function, the risk free rate, etc. In our initial calibration, we set $L = 5$ and $l=1$; as we see below, these result in expected returns close to those in the data. In order to determine whether the heterogeneous-holdings equilibrium exists, we follow the procedure in BH. The intuition underlying the solution method in BH is to search for an equilibrium with two groups of investors. One group's optimal portfolio is the market and the risk free asset, while the other's is the market, a position in the risk free asset and a long position in the jackpot stock. Given the pay-off structure of the jackpot asset, BH search for a price for that asset such that both groups have the same utility for their optimal portfolios. They find that such a price (and hence equilibrium) exists only if the jackpot security is sufficiently skewed.

For a given value of q , we follow BH and search for a price P , such that the heterogeneous holding equilibrium exists. Table 1, Panel B shows prices, returns in each state, and expected returns for equilibria with values of q ranging from 0.01 to 0.10. As the Table shows, for $q=4\%$, the gross returns in the good state are 4.16 and in the bad state are 0.83. These values of q , L and l are quite similar to those observed in the sample: 3.9%, 4.24 and 0.91, respectively. The close correspondence between the model results and the data indicates that stocks in the highest portfolio of distress risk are sufficiently lottery-like to have negative expected returns in the BH model. Using our calibrated parameters the expected excess return over the risk free rate for such stocks is -5.4%. Note that in the calibration, the market risk premium is 7.5% per year. Therefore, compared to the market, the high DEATHP portfolio

should earn a return of -12.9% per year according to the model; this return is of similar magnitude to the excess return over the market that the high DEATHP portfolio earns in the data, of -13.1% per year.⁵

These results indicate that the possibility of earning jackpot returns in our sample has the potential for explaining these securities' low subsequent returns.

3 A Logit Model for Jackpot Returns

In this section, we build a model to predict the *ex-ante* likelihood of jackpot returns, and examine whether the probability of jackpot is correlated with the probability of distress, thereby offering an explanation of the puzzling relation between default probability and subsequent returns observed in CHS. In particular, we define jackpot returns (Section 3.1), describe our model to predict the likelihood of future jackpot payoffs (Section 3.2), examine the key determinants of jackpot probabilities (3.3), investigate alternate specifications (Section 3.4) and analyze the out-of-sample forecasting power of our model (Section 3.5).

3.1 Defining Jackpots

We define jackpot returns as log returns greater than 100% over the next year. We choose to define jackpots as a binary event for several reasons. First, this corresponds to the skewed asset pay-off in Barberis and Huang (2008), which is binary. This makes it easier for us to relate our results to those in Barberis and Huang (2008), as in the calibration exercise described above. Second, similar to bankruptcy, an extremely high return is a salient event that attracts investor

⁵ We also try additional calibrations, similar to Barberis and Huang (2004) where we use the full distribution (in bins of 50 stocks) of returns and search for a value of 'c' such that when returns of high distress risk stocks are shifted by 'c' the heterogeneous holdings equilibrium obtains. We find that high distress stocks are underpriced according to the model even after taking into account their low returns: we need to reduce the returns of high distress stocks by approximately 0.6% per month for the equilibrium to exist.

attention and is easier to understand and measure as compared to moments of the return distribution. Prior research has shown that investors' risk attitudes to such rare events are very different from their attitudes to normal events. Psychological studies such as Tversky and Kahneman (1992) show that investors behave as if they overweight small probability events. Equilibrium models like Liu, Pan and Wang (2005) show that rare events require a significant risk premium and this premium helps explain the option volatility smirk in the index option market. Third, defining jackpots as binary events allows us to use the same logit model that CHS used in measuring default probabilities which makes it easier to examine the relationship between the two probability measures.⁶ Since our cut-off of annual log returns in excess of 100% is ad-hoc, we try different cutoffs in defining jackpot returns and obtain similar results in our robustness tests.

Figure 1 shows the time series of *ex post* jackpot probabilities, or the fraction of firms that realize returns greater than 100% over the next year, with NBER dated recessions in gray. The time series suggests that the chance of earning jackpot returns is typically high just when the economy is coming out of a recession, although there are exceptions, such as the 1997-1999 period when internet stocks did exceptionally well.

3.2 A Logit model to predict jackpots

We model the probability of a firm achieving a jackpot return in the next 12 months as a logistic distribution given by:

⁶ There are other possible setups. For example, we could also estimate risk-neutral skewness, as in Bakshi, Kapadia and Madan (2003) and Conrad, Dittmar and Ghysels (2012). However, this would require us to specify a pricing kernel to convert risk-neutral skewness into the physical measure, in order to assess the extent to which this measure could explain the CHS results. For comparison to CHS, we choose to use physical measures of tail events instead. In addition, there are relatively few firms that have high default probability and sufficient traded options to compute risk-neutral skewness.

$$P_{t-1}(Jackpot_{i,t} = 1) = \frac{\exp(a + b \times X_{i,t-1})}{1 + \exp(a + b \times X_{i,t-1})},$$

where $Jackpot_{i,t}$ is a dummy variable that equals one if the firm's log return in the next 12 month period is larger than 100%, and $X_{i,t-1}$ is a vector of independent variables known at $t - 1$. An increase in the value of $a + b \times X_{i,t-1}$ indicates that the probability of achieving a jackpot return in the next 12 month is higher. For each firm, we begin by estimating the parameters of a baseline logit model using at least 20 years of historical data, and then construct out-of-sample estimates of jackpot probabilities. We re-estimate this model once a year (in June), to avoid overlapping returns.

We use variables employed by prior skewness research (Chen, Hong, and Stein (2001) and Boyer, Mitton and Vorkink (2010)) to predict jackpot returns. These variables include the stock's (log) return over the last 12 months (RET12), volatility (STDEV) and skewness (SKEW) of daily log returns over the past three months, detrended stock turnover (TURN : six month volume / shares outstanding minus 18 month volume / shares outstanding), and size (SIZE: log market capitalization). We augment these variables with three new variables: firm age (AGE: number of years since first appearance on CRSP), asset tangibility (TANG: Gross Property Plant and Equipment / Total Assets), and sales growth (SALESGRTH) over the prior year. Appendix 1 provides further details on the construction of these variables. Our priors are that young, rapidly growing firms with less tangible assets are more likely to exhibit extremely high returns. All accounting data are lagged by six months, to ensure that these data are known to investors, and all independent variables are winsorized at 5% and 95%, following CHS. The accounting variables that we use to predict a jackpot return are constructed from annual CompuStat data. Our sample begins with accounting data for the year ending December 1950, matched to CRSP

data for June 1951. The data are relatively sparse initially, with 326 firms with non-missing total assets at the end of June 1951, rising to a maximum of 6834 firms in June 1997, and declining to 3972 firms at the end of our sample in 2009. Any potential survivorship bias in the early years of the CompuStat sample does not affect our results, since all of our key tests are based on out-of-sample predicted probabilities for jackpots beginning in 1972. Beginning our sample construction in 1951 enables us to have 20 years of data to estimate our first set of out-of-sample jackpot probabilities in 1972.

Table 2, Panel A provides summary statistics for these variables over the 1951-2009 sample period. Panel B examines these variables for firms that subsequently realized jackpot returns over the next year. Jackpot firms tend to be smaller, younger, more volatile, with fewer tangible assets, and have lower prior returns than firms on average. We examine the relative importance of these variables in a multivariate context below.

3.3 What predicts jackpot returns?

Table 3, Panel A reports results from our baseline model for predicting jackpot returns. All variables are statistically significant. Stocks with higher past skewness, higher returns in the past 12 months, higher sales growth rate and higher volatility are associated with a higher probability for jackpots. Note that although the results in Table 2 show that jackpot firms have low average past returns in a univariate context, this negative correlation is not evident once we control for the other variables in our model. Younger firms, firms with less tangible assets, lower stock market turnover and smaller stock market capitalization are more likely to have jackpot returns. Nevertheless, the importance of these variables is quite different. In the fourth column, we report the percentage change in the odds ratio for a one standard deviation change in the independent variable. The odds ratio is the log of the ratio of the probability of a jackpot return

divided by the probability of not achieving a jackpot return. Among all the variables, AGE, STDEV and SIZE have the largest impact on the odds ratio of the logistic regression. A one standard deviation increase in firm age reduces the odds ratio for jackpots by 27%, a one standard deviation increase in STDEV increases the odds ratio by 32.9%, and a one standard deviation increase in firm size reduces the odds ratio by 34.9%.

3.4 Alternate specifications

We explore several other specifications for forecasting jackpot returns in order to understand the robustness of our results to model specification. Panel B of Table 3 reports alternative logit models. In Model 2 we replace the size variable in the baseline model with two dummy variables: *SMALLDUMMY* and *MEDIANDUMMY*. *SMALLDUMMY* is a dummy variable which equals one if the market capitalization of the stock belongs to the bottom tercile and equals zero otherwise. Similarly, *MEDIANDUMMY* is a dummy variable for the middle market capitalization tercile. The introduction of the size dummy variable is to accommodate potential non-linearity in the size effect. We also add in a dummy variable for firms listed on NASDAQ. The two size dummy variables and the NASDAQ dummy all show up significantly.

Model 3 includes exactly the same variables used in CHS. Variables are constructed as described in their paper, except that ours are based on annual rather than quarterly CompuStat data (see Appendix 1 for details). These variables are return on market assets (ROMA), relative size (RELSIZE), market leverage (MLEV), cash and short-term investments as percentage of market equity and total liability (CASH), average past 12-month return over S&P500 index return (EXRAVG), return volatility (STDEV), market-to-book ratio (MB) and the log of stock price truncated at \$15 (PRC15). All enter the logit regression with a significant coefficient with the exception of ROMA and MLEV. RELSIZE and PRC15 are inversely related to jackpots as

well as to default. That is, being a relatively small size firm or low priced stock increase the probability of both jackpot returns and distress, consistent with a limits-to-arbitrage explanation of why these returns persist. High cash holding reduces the distress probability but increases jackpot probability; high past EXRAVG lowers a firm's distress probability and increases jackpot probability. STDEV has the same effect on both distress and jackpots, with high STDEV associated both with high distress and jackpot probabilities. Being a growth firm (high MB) also increases both distress and jackpot probabilities. Return on market assets (ROMA) and market leverage (MLEV) are significant in predicting distress but are not significant in forecasting jackpots. Both model 2 and model 3 achieve higher pseudo R-squares of 6.07% and 6.39%, respectively. We show in subsequent tests that these alternative models produce similar results in out-of-sample forecasts as our baseline model, and that returns earned in strategies based on these models are also similar to those obtained using the baseline model.

3.5 Predictive power

The baseline logit model achieves a pseudo R-square of 5.76%. The relatively low R-square is not surprising, since it is well-known that extreme events are difficult to forecast. We test whether this relatively low predictive power allows us to generate reliable measures of jackpot returns out-of-sample. Starting from 1951, we use all available data (expanding annual rolling windows) to re-estimate our baseline model and then generate out-of-sample forecasts for the probability of jackpot returns with each set of estimated parameters. The first predicted out-of-sample jackpot return is in 1972; the last out-of-sample forecast is in 2009. The 1972-2009 forecasts match the availability of the CHS default probability measure that we use, along with the jackpot probabilities reported in other tests. We borrow from the default prediction literature (see Vassalou and Xing (2004)) in using the accuracy ratio to evaluate the effectiveness of the

out-of-sample predictability. The accuracy ratio reveals the ability of a model to predict actual jackpot returns over a one-year horizon. A completely uninformative model will yield an accuracy ratio of zero and a perfect model will yield an accuracy ratio of 100%. Appendix 2 describes the construction and rationale underlying the accuracy ratio.

The out-of-sample predicted jackpot probability from our baseline model has an accuracy ratio of 77.41% in predicting realized jackpot returns. In particular, the 63% of stocks that realize an *ex post* jackpot return are in the top 1% of *ex-ante* predicted jackpot probability; 70% of stocks that realize a jackpot return are in the top 10% of predicted jackpot. We also compare the predictability of our model to the predictive power obtained using only volatility to predict jackpots. Even though volatility is very important in forecasting jackpots, our out-of-sample predicted jackpot probability measure has substantially higher predictability for realized jackpots than volatility does. Using volatility alone, only 3.36% of stocks that realize jackpot returns are in the top 1% of stocks with highest volatility; 23.88% of stocks that realize jackpot returns are in the top 10% of volatility. The accuracy ratio for volatility itself is only 37.8%. This shows that our model is not just driven by volatility – other variables also matter a great deal in predicting jackpots.

Foreshadowing our primary result, we compute the accuracy ratio from using DEATHP, the probability of default according to the model in CHS, to predict jackpots. We find that DEATHP does remarkably well, with an accuracy ratio of 54.07%. Thus, predicted default does a much better job of predicting jackpots than using volatility alone, although it fares worse than our full model for predicting jackpots out-of-sample. This result also indicates that the commonality between the two measures of predicted distress and jackpots is not driven by volatility alone.

4 Can the probability of jackpot returns explain the distress risk puzzle?

In this section, we test whether a high probability of earning jackpot returns can explain the low average returns of high distress risk stocks. First, in Section 4.1, we examine whether stocks with a high probability of jackpot returns have low average returns. In Section 4.2, we investigate the ownership structure of stocks sorted on the basis of DEATHP and JACKPOT. In Section 4.3 we examine the effect of limits to arbitrage on the returns of stocks with high jackpot probability. In Section 4.4, we compare both characteristics and factor loadings of distress sorted portfolios and predicted jackpot sorted portfolios. Finally in Section 4.5 we analyze the correlation between the probability of jackpots and the probability of distress and examine how the distress strategy and jackpot strategy are related to each other.

4.1 Average returns for strategies based on predicted jackpot probability

We examine whether trading strategies based on the probability of jackpot returns can generate similar return patterns as those based on CHS default probability. At month t , we use out-of-sample predicted jackpot probability computed using available information to sort all stocks into ten deciles and compute value-weighted portfolio returns for month $t+2$. Note that we skip a month between portfolio formation and measuring returns to alleviate concerns regarding the potentially confounding microstructure effects such as bid-ask bounce. Portfolios are rebalanced each month.⁷ As described above, our out-of-sample predicted jackpot probability measures begin in 1972, to allow at least 20 years of data for the initial estimation.

⁷ The rebalancing is due to changes in market variables like size, volatility and past annual returns, since the other accounting based variables only change annually. We have similar results if we rebalance our portfolios annually (four factor alpha for the '1-10' portfolio of 0.65% per month) or if we do not skip a month between measuring jackpot probability and returns (four factor alpha for the '1-10' portfolio of 1.1% per month). These results are available upon request.

In Table 4, we report the results from tests on value-weighted decile portfolios formed from sorts on out-of-sample predicted jackpot probability. In Panel A, we report average excess returns over the risk free rate for these portfolios as well as the alphas estimated from three different models: CAPM, Fama and French (1993) three-factor model, and Carhart (1997) four-factor model. The average excess returns in the first row of Panel A do not show a monotonic pattern. In fact, average excess returns increase from decile one to decile four before decreasing. The sharp drop in excess returns comes in decile nine (0.03% per month) and decile ten (-0.62% per month). A long-short portfolio that holds the decile of stocks with the lowest jackpot probability and goes short the decile with the highest jackpot probability yields an average return of 1.06% per month.

Turning to risk-adjusted returns, we find that controlling for risk using the CAPM, Fama and French three-factor or Carhart (1997) four-factor model does not help explain the low returns of the portfolios with high jackpot probability. In fact, if anything, the poor performance of high predicted jackpot probability stocks looks worse after using these models. The alpha on the long-short portfolio increases to 1.39% for the CAPM, 1.39% for the Fama French three-factor model, and 1.1% for the Carhart (1997) four-factor model respectively. In each model, the alpha is highly significant. In Panel B of Table 4, we report the loadings on MKT, SMB, HML and WML in the four-factor model for the ten jackpot portfolios. The variation in factor loadings across these portfolios is striking. The loading on MKT increases from decile 1 to decile 9 and then falls slightly in decile 10. The SMB loading across the ten jackpot portfolios increases monotonically from -0.26 in decile 1 to 1.62 in decile 10 and the HML loading decreases from 0.04 to -0.48 going from lowest jackpot probability decile to highest jackpot probability decile. This is indicative of the prevalence of small and growth stocks in the deciles with high jackpot

probabilities. High jackpot probability stocks are also likely to be loser stocks as they load negatively on the momentum factor WML.

The jackpot strategy (going long the least likely and short the most likely decile of JACKPOTP) has an annualized excess return of 12.71% and a standard deviation of 25.88%. The Sharpe ratio is 0.49, higher than the stock market (0.32) over the same time period, and comparable to that of HML (0.49) and WML (0.55). The first eight deciles of jackpot sorted portfolios have negatively skewed portfolio returns, while deciles 9 and 10 have positive skewness of 0.13 and 0.70, respectively. In untabulated results, we find that over our sample of 1972 to 2009, the long-short jackpot strategy is in fact the best-performing strategy among MKT, SMB, and HML in terms of cumulative returns.

4.2 Institutional Ownership

In the previous section, we present evidence that investors bid up the price of securities with a high probability of a jackpot payoff; investors exhibit a preference for skewness. BH suggest that individual investors are more likely than institutions to display a preference for stocks with lottery-like payoffs. Kumar (2009) finds that retail investors exhibit a preference for stocks with lottery-like features, while institutions do not. We therefore investigate the ownership structure of stocks sorted on the basis of DEATHP and JACKPOTP. Institutional ownership is defined as the fraction of shares owned by institutions in the Thomson Reuters Institutional Holdings database. Figure 2 plots the median institutional ownership for the median firm and for the tenth deciles of DEATHP (Panel A) and JACKPOTP (Panel B). Consistent with prior research (e.g. Bennet, Sias, and Starks (2003)), Panel A shows that median institutional ownership has increased dramatically in the United States from 15% at the start of the sample in 1980 to 58% at the end of the sample in 2009. High DEATHP stocks have substantially less

institutional ownership than the median stock for which DEATHP data is available, with their institutional ownership increasing from 3% at the start of the sample to 15% at the end. The difference between institutional ownership for the median stock for which DEATHP data is available and a high DEATHP stock has increased over time: a regression of this difference on a linear time-trend results in a significant coefficient on the time-trend of 0.68% per year (not tabulated). Thus, the increase of institutional ownership for the median stock has been more rapid than that for a stock in the highest DEATHP quintile. Similarly Panel B shows that high JACKPOTP stocks also have lower institutional ownership than the median stock (for which data is available to compute JACKPOTP).⁸ The difference in institutional ownership between the median firm and high JACKPOTP firms also increases significantly over time, with a coefficient of 1.05% per year.

We also find evidence that institutional ownership is sensitive to increases in default and jackpot probabilities. Panels C and D present results of an ‘event-study’ where the event is the entry of the firm into the top DEATHP (Panel C) or JACKPOTP (Panel D) quintile. We de-mean institutional ownership by subtracting the mean institutional ownership for all firms that month, and compute average demeaned institutional ownership for all firms in event time. Panel C shows that demeaned institutional ownership declines from -12.5%, four quarters prior to the entry into the top decile of DEATHP, to -16.3% at the quarter of entry into the top decile, and declining further post-entry to -17.8% four quarters after entry into the top decile. Note that both changes (from Quarter -4 to Quarter 0 and from Quarter 0 to Quarter 4) are statistically

⁸ Note that the median firm is defined differently in the two panels. In Panel A, we compute the median over all firms for which DEATHP is available, while in Panel B we do it for all firms for which JACKPOTP is calculated. Since CHS substitute cross-sectional means for any missing data items, while we do not include firms with missing data while computing JACKPOTP, DEATHP is available for more firms than JACKPOTP.

significant with p-values from a test of equality of means of less than 0.01. Similarly, Panel D shows that de-meaned institutional ownership around entry into the highest JACKPOTP decile declines from -16.5% to -18.4% from quarter -4 to quarter 0 and further declines to -19.8% in quarter +4. Again, both changes are statistically significant, with p-values of less than 0.01.⁹

Overall, this section shows that despite the increase in institutional ownership in the US stock market, stocks classified as high DEATHP or high JACKPOTP firms have a majority of stocks held by retail investors. Also, the increase in institutional ownership for such firms over time has been slower than that of the median firm. Finally, institutional ownership declines significantly four quarters prior and four quarters subsequent to entry into the top DEATH or JACKPOT decile. This suggests that retail investors ('non-institutions') are net buyers of such stocks as they approach default or become more likely to be jackpots.

4.3 Limits to arbitrage and the Jackpot effect

In the BH model, which is populated solely by investors with prospect theory utility functions, the high prices and low expected returns for jackpot assets are not arbitrage opportunities, since they are consistent with preferences in equilibrium. However, BH argue that in a more realistic setting, where both expected utility and prospect theory utility based investors coexist, limits to arbitrage may result in expected utility investors being unable or unwilling to short-sell jackpot assets in order to exploit their low returns. We therefore test the hypothesis that the jackpot effect is stronger where limits to arbitrage are high, using three measures of limits to arbitrage: size, institutional ownership and analyst coverage. Following CHS, for each limits to

⁹ The significant changes in institutional ownership in the four quarters *prior* to the entry of the firm into the top DEATHP or JACKPOTP decile is consistent with investors rebalancing based on *ex ante* estimates of default and jackpot payoff probabilities.

arbitrage variable, we first sort stocks into two groups, one with low and the other with high levels of the variable every month. Then within each group, we form portfolios long the 20% of stocks with highest jackpot probability and short the bottom 20%. Each of these portfolios is value-weighted and skips a month between portfolio formation and measuring returns. Table 5 presents mean returns, CAPM and four-factor alphas for each of these long-short portfolios.

Our first variable is size. Small firms are defined as firms smaller than the NYSE 30% cut-off, while large firms are larger than the NYSE 70% cut-off. It is clear that the jackpot effect is strong in small firms and not present at all in large firms. The long-short portfolio has returns of -1.18% (t-value -3.95) per month amongst small firms and average returns of -0.05% (t-value = -0.19) per month amongst large firms. Similarly four factor alphas are at -0.86% (t-value = -2.96) per month for small firms and at 0.01% (t-value 0.05) per month for large firms.

Our next two variables, residual institutional ownership and residual analyst coverage explicitly control for size, so that results for these variables are not just restatements of the results for size (since raw values of these variables are highly correlated with size). To control for size, we follow CHS and compute residuals from a regression of each of these variables on RELSIZE and time dummies. We then sort residuals into terciles and examine the long-short jackpot portfolio within each tercile. Institutional ownership is the fraction of shares owned by institutions in each stock from the Thomson-Reuters dataset on WRDS, and is available quarterly. It is particularly interesting to examine institutional ownership in the context of the jackpot effect. First, both the results of the preceding section, and the evidence in Kumar (2009) indicates that individuals and institutions have different preferences for lottery-like stocks; individuals display a strong preference for such stocks, while institutions are relatively averse to these securities. Second, institutional ownership is a commonly used proxy for the supply of lendable shares in

the short-selling market (see Nagel (2005) for example). Since arbitragers must short stocks with high jackpot probabilities to take advantage of their abnormally low returns, a low supply of shortable shares is an important limit to arbitrage.

We find that the jackpot effect is much stronger in firms with low residual institutional ownership, with returns of the long-short portfolio of -1.23% (t-value = -2.76) per month versus -0.33% (t-value = -0.86) for the high residual institutional ownership sub-sample. Also, four factor alphas of -1.19% (t-value = -4.45) for the low institutional ownership sub-sample are much larger in absolute magnitude than those of -0.41% (t-value = -1.59) for the high institutional ownership sub-sample. It is also interesting to note that the size control works; both the low and the high residual institutional ownership samples have similar average size and spread in jackpot probability between the long-short portfolio in each institutional ownership groups. This shows that stocks with similar jackpot probability have very different average returns, depending on the level of limits to arbitrage.

We find broadly similar results, although with smaller magnitudes, for residual analyst coverage. Analyst coverage is the natural log of one plus the number of analysts that have issued an earnings forecast for the stock on the IBES dataset within the last fiscal year. The jackpot effect is stronger for firms with low residual analyst coverage, with four factor alphas equal to -0.54% per month (t-value = -1.52) as compared with -0.25% (t-value = -0.79) for firms with high residual analyst coverage. As with residual institutional ownership, both low and high analyst coverage firms have similar spreads in jackpot probabilities and average size.

To summarize, these results demonstrate that the jackpot effect is concentrated amongst stocks with high limits to arbitrage. As a consequence, high limits to arbitrage may help explain why these pricing effects persist in the data.

4.4 The similarities and differences between high predicted jackpot probability and high predicted distress firms

Table 6 presents characteristics of firms that are in portfolios formed from sorts on DEATHP, the default probability measure in CHS, and those formed from sorts on JACKPOTP, our out-of-sample predicted jackpot measure. First, it is interesting to note that the fraction of firms in the top decile portfolio that subsequently realize a jackpot return is slightly higher (3.9%) for the highest DEATHP portfolio than for the highest JACKPOTP sorted one (3.6%). Second, as both DEATHP and JACKPOTP increase, size and past twelve month returns decrease and M/B ratios increase, although the magnitudes are different, especially for past twelve month returns. Third, DEATHP portfolios display no pattern in sales growth, increasing leverage and declining skewness of daily log returns. JACKPOTP portfolios display very different patterns for these variables, with increasing sales growth, flat leverage and increasing skewness. Note that the skewness reported in this table is the skewness in daily log returns over the *past* three months, while the skewness reported in Table 1 is the skewness of returns over the *next* year. Also, although higher default risk portfolios display declining average skewness, there is dispersion in skewness within each portfolio. We show later (in Section 5.1.1) that the low returns of distressed stocks are only present for distressed stocks with high past skewness.

Figure 2 also shows that DEATHP and JACKPOTP sorted portfolios have similar patterns in factor loadings for MKT, WML, and SMB, but have sharply different patterns for HML. Loadings on HML increase as DEATHP increases, but decrease as JACKPOTP increases. This is surprising as both sets of portfolios have similar patterns in B/M ratios. This difference in loadings is related to differences in leverage: when we restrict stocks in the highest default probability portfolios to have smaller leverage, HML loadings decline (in untabulated results).

We return to these differences in characteristics in section 5.1, where we use them to try to distinguish between distress and jackpots. In the next section, we examine whether the similarities described above result in significant correlations between the probability of default and the probability of jackpots, as well as whether returns on a trading strategy that exploits the jackpot effect are correlated with those of a strategy that exploits the default effect.

4.5 The relation between distress and jackpots

For jackpot returns to be a plausible explanation for the low returns of high distress stocks, *ex-ante* measures of these two variables should be significantly correlated with each other. In this section we investigate the relation between *ex-ante* distress and jackpot probabilities. Table 7, Panel A presents pairwise Spearman correlations between predicted distress from the CHS model (DEATHP) and different measures of the out-of-sample probability of a jackpot return. JACKPOTP, the predicted probability of a jackpot return from our baseline model, has a correlation of 41.8% with the probability of distress. We also examine the alternate prediction models from Section 3.4 that use different variables, and three additional models that use different cut-offs in defining jackpots. For jackpot predictions from alternate models and other jackpot cut-offs, pair-wise correlations with distress are in the neighborhood of 40%. Thus, firms with a high potential for ‘death’ are also likely to have a high potential for a jackpot payoff.

Next, we examine the correlation between returns of a long-short strategy designed to exploit the CHS distress effect and one designed to exploit the jackpot effect. The distress strategy, DEATHPLS, is long stocks in the bottom DEATHP decile and short stocks in the top DEATHP decile, while the jackpot strategy, JACKPOTLS, is long stocks in the bottom decile of JACKPOTP and short stocks in the top decile JACKPOTP. All portfolios are value-weighted.

Going long the safest stocks or those least likely to achieve a jackpot return ensures positive average returns for both strategies.

In Panels B and C of Table 7, we examine the correlation in the returns of the jackpot and distress strategies, and compare their exposures to the four standard factors. The first specification in Panels B and C reports how returns of the two strategies co-vary with one another. The results indicate a strong relation, with 32.5% of the time-series variation in the jackpot (distress) strategy return explained by the distress (jackpot) strategy return. In both specifications, the alpha estimates decline sharply and become statistically insignificant. For example, recall from Table 1 that the four-factor alpha of the distress strategy is 1.21% per month; the alpha when jackpot returns are included in the regression is 0.54% per month.¹⁰

In the right-hand-column of Panels B and C, we include other risk factors (MKT, SMB, HML and WML) in the analysis. Controlling for these risks leaves both the distress and the jackpot strategy with significant alphas, of 0.54% and 0.4% per month, respectively. That is, by design, both distress and jackpot strategies appear to be relatively low risk investments.

We do not interpret these regressions as a formal asset pricing model. Rather, these results indicate that there is a significant relation between distress and jackpot strategies; the returns of portfolios sorted on the probability of jackpot returns are correlated with those sorted on the probability of distress. Thus, a high probability of a jackpot return is a plausible

¹⁰Chava and Purnanandam (2010) find that the distress effect documented in CHS is only statistically significant in the 1980's. In a longer sample, with the same model for distress as CHS, we find the distress effect is significant over other subsamples from 1972 to 2009, although it is largest in the 1980s. The difference in our results is likely due to differences in the models used to forecast jackpot. Chava and Purnanandam (2010) use the same model as Shumway (2001), while CHS modify that model and show that their modifications improve forecasting power for defaults. For example, for their definition of profitability, CHS use geometrically declining weights over the past four quarters of Net Income / Market Total Assets, while Chava and Purnanandam (2010) using annual Net Income / Book Total Assets.

explanation for the low average returns of stocks with high default probability although, as the factor loadings in Figure 2 indicate, there are some differences between the strategies' risk profile. In the next section, we try to use these differences to disentangle the relation between jackpots and distress further.

5 The relation between distress and jackpots effects on expected returns

This section investigates the relation between the effects of distress and jackpots on expected returns. In Section 5.1, we use firm characteristics, motivated by economic arguments, to construct subsamples of firms that have similar probabilities of distress but different probabilities of jackpots.¹¹ These allow us to investigate whether the effect of high default risk on returns is present when jackpots are relatively unlikely. In Section 5.2 we use Fama-MacBeth regressions to test whether one effect subsumes the other.

5.1 Interesting sub-samples

We use the following characteristics: past return skewness, and sales growth and M/B to construct sub-samples where although the level of default risk is high, the probability of jackpots is low. We have strong prior beliefs and empirical evidence (in Section 3) that these characteristics predict jackpot returns (but not defaults). Note that for ease of comparison, all

¹¹ We attempted to create a sub-sample of firms with low default probabilities and varying jackpot probabilities, in order to investigate whether the jackpot effect is concentrated in firms with high default probabilities. Specifically, in a previous version of the paper we report that in the subsample of firms with zero leverage, which might be associated with low default probabilities, the jackpot strategy provides significant returns (a four factor alpha of 0.58% per month after skipping a month, and 1.3% without skipping a month). Thus, the jackpot effect is present even for firms with no debt; however, even firms with no leverage may have significant probabilities of economic distress. And, as stated previously, attempts to use double-sorting on jackpot and distress probabilities resulted in samples that were too small to draw reliable inferences.

results in this section are for the sample where both probabilities of distress and probabilities of jackpot returns can be computed.

5.1.1 Test 1: Skewness

We first sort firms in the highest CHS default probability quintile based on the skewness of their daily log returns over the past three months. As before, portfolios are value-weighted and we skip a month between portfolio formation and measuring returns. Note that the difference in default probabilities between low-skewness and high-skewness firms is modest (2.28% vs. 2.40%), with a relatively larger difference in jackpot probabilities across the low-skew and high-skew subsamples (1.20% vs. 1.48%), as expected. Table 8 shows that amongst firms in the highest quintile of default probability, firms with skewness less than the 30th percentile for the full sample do not have abnormally low four factor alphas (-0.21% per month). Instead, the low returns of high default probability stocks are concentrated in the portfolio of firms in the top 30% of skewness, with a four-factor alpha of -0.75% per month. The difference in subsequent returns between these two portfolios is statistically significant, with a four factor alpha of 0.54% per month (and a t-value of 2.21) for the portfolio which is long the low skewness firms and short the high skewness firms amongst the subsample of firms with high default probability.

5.1.2 Test 2: Sales growth and M/B

We sort firms in the highest CHS predicted default probability quintile into ‘traditional distressed’ firms and ‘speculative’ firms. Traditional distressed firms are defined as those firms that have low sales growth and low market-to-book ratios (in the bottom 30% of the full sample for both), while speculative firms are those with high market-to-book ratios and high sales growth (in the top 30% of the full sample for both). This procedure allows us to create sets of portfolios with similar distress probabilities but different jackpot probabilities.

Panel B of Table 8 presents results for value-weighted portfolios of both traditional distress firms and speculative firms. Within the quintile of the highest distress probability, on average, 169 stocks are traditional distressed firms, while 53 stocks are categorized as speculative. The average annualized default probability for traditional distress stocks (2.64%) is, if anything, higher than that for speculative distress stocks (2.28%). However, traditional distressed firms do not earn abnormally low returns in subsequent periods, and have a four-factor alpha of -0.14% that is not significantly different from zero. In sharp contrast, the *speculative* firms in the highest CHS default probability quintile have a four-factor alpha of -1.70% per month with a t-value of -4.81. Even though the average distress probability is similar between these two groups, the average probability of a jackpot return is much higher for the speculative stocks at 1.97% as compared to 1.09% for traditional distress stocks. Thus, the results in Panel B of Table 8 show that among stocks with high default probability, traditional distressed firms do not go on to earn abnormally low returns, while speculative firms do. Both sets of firms have similar default probabilities, but speculative firms have twice the jackpot probability of traditional distressed firms.

This test is reminiscent of previous results in the literature on the effect of distress risk on the cross section of stock returns. Dichev (1998) shows that stocks classified as high default risk according to the Ohlson (1980) model have low average returns. Griffin and Lemmon (2002) show that these low returns are concentrated in stocks with high M/B ratios. They argue that such stocks are overvalued. We show that the low returns that CHS report are concentrated in speculative stocks as described above and not in traditional distressed stocks. We also show that such speculative stocks have much larger probabilities of jackpot returns than traditional distressed stocks, providing an economic rationale for their high valuations. Overall, our results

are consistent with the hypothesis that, holding the effect of distress risk constant in the sample, returns to stocks are significantly lower in stocks that have high jackpot probabilities.¹² The results suggest that, on average, it is not merely a high default probability that leads to investors bidding up the price of a security—the payoff distribution must have a significant probability of a relatively large positive outcome as well, in order to become ‘overpriced’.

5.2 Fama-MacBeth regressions

Fama-MacBeth regressions allow us to test whether default or jackpot probabilities are important in predicting returns while controlling simultaneously for other variables (like past volatility) that have been shown to affect expected return and are correlated with both these measures. For our Fama-MacBeth regressions, we standardize each explanatory variable to have a mean of zero and a standard deviation of one in every cross-section, by subtracting the variable’s cross-sectional mean and dividing by its cross-sectional standard deviation. This allows coefficients to be easily interpretable: each coefficient measures the effect on returns of a one standard deviation shock to the explanatory variable. We also lag each explanatory variable by one month—skipping a month between measuring the explanatory variable and returns allows us to control for microstructure-related effects such as bid-ask bounce.

In Table 9, specification 1, we include the following variables that have been shown to have an effect on returns in the cross-section: book-to-market (BM), lagged past 12 month return ($\text{Ret}(t-12,t-2)$), market capitalization (SIZE) and log stock return volatility (LOGSTDEV). All

¹² An earlier version of this paper also reported results for independent portfolio sorts on distress and jackpot probabilities. We find broadly similar results with these portfolios as those with Fama MacBeth regressions and do not report them in this version for brevity. Also, since the extreme high distress and low jackpot probability, as well as the low distress and high jackpot probability portfolios are sparsely populated, portfolio results are less reliable than the Fama-MacBeth results we report below.

variables are observed at the end of the previous month. All variables carry significant coefficients with the expected sign. Specifically, coefficients on BM (positive), $\text{Ret}(t-12,t-2)$ (positive) and size (negative) are consistent with stylized facts concerning value, momentum and size effects. In specification 2, we add predicted default probability, DEATHP, to these firm characteristics. The coefficient on DEATHP is -0.42% and is statistically significant. In specification 3, we drop default probability and add jackpot probability, JACKPOTP. The coefficient on JACKPOTP at -0.37% per month is very close to that of DEATHP and is statistically significant. When both JACKPOTP and DEATHP are introduced into the Fama-MacBeth regression in specification 4, we see that the sign and magnitude of the coefficients on JACKPOTP and DEATHP are very similar to previous specifications. Also, it is interesting to note that volatility is not significant in the final specification. These results suggest that the results in Ang, Hodrick, Xing and Zhang (2008) which show that idiosyncratic volatility is negatively related to future returns, is linked to the possibility of death and jackpot returns. These results are consistent with Boyer, Mitton, and Vorkink (2010) and Bali, Cakici, and Whitelaw (2011) who find that a potential for skewness or lottery-like pay-offs explains the idiosyncratic volatility effect.

In unreported tests, we find that these results are robust to alternate transformations of variables. For example, using log default or jackpot probabilities, not standardizing variables, or not logging volatility produce similar results. Overall, this section shows that controlling for distress does not drive out the effect of jackpots on expected returns, nor does controlling for a high probability of a jackpot return drive out the effect of distress on expected returns. If

anything, the Fama-MacBeth regressions suggest that *both* distress and jackpot probability have an impact on expected returns.¹³

6 Robustness tests

In Table 10 we examine the robustness of our results with the jackpot trading strategy. We report the jackpot strategy's performance 1) with different specifications to forecast jackpot returns; 2) in each decade separately; and 3) with different cut-offs in the definition of a jackpot return. All portfolios formed continue to be value-weighted. For brevity, we only report the alpha from the Carhart (1997) four-factor model in this table.

In Panel A, we use the alternate specifications discussed in Section 3.4 to forecast the probability of a jackpot return. Specifically, in Model 2, we introduce nonlinearity in size by adding two indicator variables for size categories. In addition, we add a dummy variable for NASDAQ firms. With this specification, we first re-estimate the out-of-sample jackpot probability and then re-construct the jackpot strategy. We skip a month before portfolio formation to avoid short-term return reversal effect. The one minus ten portfolio return has an alpha of 0.87% per month with a t-statistic of 3.58. Model 3 includes exactly the same set of

¹³ The Fama-MacBeth cross-sectional regression results should be interpreted with some caution. Given the high correlation between the probability of distress and the probability of a jackpot return, which we observe in our sample, it is possible that collinearity affects these results. Related, measurement error in the predicted probabilities of distress and jackpot returns may make it difficult to distinguish between the two variables in Fama-MacBeth regressions. We investigate this possibility by performing a simulation exercise using Fama-MacBeth regressions with correlated independent variables in Appendix 3. In our simulation, expected returns depend only on the probability of earning a jackpot return, but the probability of jackpots and death are correlated. In addition, the probabilities of jackpots and death are measured with error. When observed values of the probability of death and jackpot returns are used as explanatory variables in Fama-MacBeth regressions, we find that both are statistically significant, across a range of plausible values for measurement error and true correlation. These results suggest that given the observed correlation between death and jackpot returns, the results of Fama-MacBeth regressions may not allow us to determine whether one variable is more important than another in driving the effect in returns.

variables as in the CHS distress forecasting specification. Both model 2 and model 3 generate very similar jackpot strategy returns as our baseline model. There is a sharp drop in the four-factor alpha in decile 10 for both models. Overall, regardless of the specification used, the jackpot strategy returns remain statistically and economically significant.

In Panel B of Table 10, we report results from our baseline model for different sub-periods. We divide our full sample into four sub-samples: 1972 to 1979, 1980 to 1989, 1990 to 1999 and 2000 to 2009. The jackpot strategy has the highest return in the 1980-1990 sub-period, with a four factor alpha of 1.67%. Though smaller in magnitude, in other sub-samples, the jackpot strategy return remains strongly positive and significant at conventional levels in all sub-samples with the exception of 1990-1999 subsample, where the p-value on the estimated alpha of 0.89% per month is 0.08.

Next we examine whether the low average returns of stocks with a high probability of earning a jackpot return depend on how jackpot returns are defined. In our baseline model, jackpot returns are defined as log returns greater than 100% over the next 12 month period. In Panel C of Table 10, we use the same baseline model to forecast jackpots out-of-sample, with a jackpot return defined as an arithmetic return above 50%, 75%, or 100% over the next 12-month period. The results are very similar to Panel A in Table 3 where a jackpot return is defined as log return above 100%. The jackpot strategy alpha increases slightly when the cutoff return to be considered a jackpot increases. All strategies are highly significant.

Summarizing, these results indicate that stocks with a high predicted probability of a jackpot return also have low subsequent average returns; this result is robust to different definitions of, and models for predicting, jackpot returns.

7 Conclusion

The CHS result that firms with high default risk earn low returns is surprising because of the direction of the effect that ‘risk’ has on average returns. We show that there is a large overlap between stocks classified as high default risk by the CHS model and those that are likely to produce extremely high returns (over 100%) over the next year. Thus, we show that these stocks possess a feature that investors desire: lottery-like payoffs that lead to high valuations and low expected returns. This is consistent with the model in Barberis and Huang (2008), where investors with prospect theory based utility functions display a strong preference for such stocks, resulting in low average returns in equilibrium.

We build a model to predict which stocks have lottery-like returns (‘jackpot’ payoffs), using a logit model similar to the model in CHS, except that our dependent variable is 1 if returns over the next year are over 100%. We estimate this model on an expanding out-of-sample window and find that stocks with a high predicted probability of jackpot returns do indeed have low average subsequent returns. We show that stocks with high default and jackpot probabilities have relatively low institutional ownership, consistent with the hypothesis that these are largely owned by retail investors who are more likely to display a preference for total skewness. We also find that low average returns are concentrated in stocks with high limits to arbitrage, consistent with the hypothesis that these limits make it difficult for investors with expected utility preferences to arbitrage these low returns away. We also find that the predicted probability of a jackpot return is highly correlated with the probability of default from the model of CHS. Therefore, a high probability of a jackpot return is a plausible explanation for the results in CHS.

We run a set of tests to examine the relation between the effects of jackpot payoffs and default on expected stocks’ returns. The returns of jackpot and distress strategy returns are

significantly correlated with each other. We use different subsamples to attempt to disentangle these effects. First, we find that amongst stocks with high default probability, low average returns are only present in stocks in the top 30% of daily return skewness in the past three months and not in those with the lowest 30% of skewness. Second, we show that the low returns that CHS report are concentrated in ‘speculative stocks’ with high sales growth and high M/B ratios and not in ‘traditional distressed’ stocks with low sales growth and low M/B ratios. These results are consistent with the interpretation that it is skewness that leads investors to hold ‘overpriced’ distressed stocks. We also use Fama-MacBeth regressions, to control for variation in both distress and jackpot probabilities, as well as other variables such as volatility. We find evidence that default probabilities do not drive out the significance of jackpot probabilities in explaining subsequent returns; however, neither do jackpot probabilities completely subsume the predictive power of default probabilities for subsequent returns.

Overall, our results suggest that a high probability of jackpot payoffs is responsible for at least a portion of the low average returns of stocks with high default probability, with investors bidding up the price of securities with high default probabilities if the securities offer a lottery-like payoff structure. In addition, the magnitude of the average abnormal return in jackpot securities varies with factors that are associated with the costs of arbitrage in these securities.

References

- Altman, Edward I., 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance* 23, 589–609.
- Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaozan Zhang, 2006, The cross section of volatility and expected returns, *Journal of Finance* 61, 259-299.
- Bali, Turan G., Nusret Cakici, Robert. F. Whitelaw, 2011, Maxing out: Stocks as lotteries and the cross-section of expected returns, *Journal of Financial Economics*, 99(2), 427–446
- Barberis, Nicholas, and Ming Huang, 2004, Stocks as Lotteries: The Implications of Probability Weighting for Security Prices, Working Paper.
- Barberis, Nicholas, and Ming Huang, 2008, Stocks as Lotteries: The Implications of Probability Weighting for Security Prices, *American Economic Review*, 98, 5, 2066-2100.
- Bennet, James, Richard Sias, and Laura Starts, 2003, Greener pastures and the impact of dynamic institutional preferences, *Review of Financial Studies*, 16, 1203-1238.
- Boyer, Brian, Todd Mitton and Keith Vorkink, 2010, Expected Idiosyncratic Skewness, *Review of Financial Studies*, 23, 169-202.
- Boyer, Brian, and Keith Vorkink, 2011, Stock Options as Lotteries, Working paper, Brigham Young University.
- Brunnermeier, M.K., C. Gollier, and J.A. Parker, 2007, Optimal beliefs, asset prices, and the preference for skewed returns, *The American Economic Review* 97, 159–165.
- Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In Search of Distress Risk, *Journal of Finance* 63, 2899-2939.
- Carhart, Mark, 1997, On Persistence in Mutual Fund Performance, *Journal of Finance*, 52 (1), 57-82.

- Chan, K., Chen, N., 1991. Structural and Return Characteristics of Small and Large Firms. *Journal of Finance* 46, 1467–1484.
- Chava, Sudheer and Robert A. Jarrow, 2004, Bankruptcy prediction with industry effects, *Review of Finance* 8, 537—569.
- Chava, Sudheer and Amiyatosh Purnanandam, 2010, Is Default Risk Negatively Related to Stock Returns? *Review of Financial Studies* 23, 6, 2523-2559.
- Chen, J., H. Hong, and J. Stein. 2001, Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices, *Journal of Financial Economics* 61, 345–81.
- Conrad, Jennifer, Eric Ghysels, and Robert Dittmar, 2012, Ex Ante Skewness and Expected Stock Returns, forthcoming, *Journal of Finance*.
- Dichev, Ilia D., 1998, Is the Risk of Bankruptcy a Systematic Risk?, *Journal of Finance* 53(3), 1131-1147.
- 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.
- Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.
- Garlappi, Lorenzo, and Hong Yan, 2011, Financial Distress and the Cross-section of Equity Returns, *Journal of Finance*, 66, 3, 789-822.
- George, Thomas J., and Chuan-Yang Hwang, 2010, A resolution of the distress risk and leverage puzzles in the cross section of stock returns, *Journal of Financial Economics*, 96, 56–79.
- Griffin, John M. and Michael L. Lemmon, 2002, Book-to-market equity, distress risk, and stock returns, *Journal of Finance* 57, 2317-2336.
- Gao, Pengje, Christopher Parsons, and Jianfeng Shen, 2012, The global relation between financial distress and equity returns, Working paper.

- Kapadia, Nishad, 2011, Tracking down distress risk, *Journal of Financial Economics* 61, 345–81.
- Kumar, Alok, 2009, Who gambles in the stock market? *Journal of Finance*, 64 (4), 1889-1933, 2009.
- Lau, H. S., J. R. Wingender and A. H. Lau, 1989, On Estimating Skewness in Stock Returns, *Management Science*, 35, 1139-1142.
- Liu, Jun, Jun Pan and Tan Wang, 2005, An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks, *Review of Financial Studies* 18, 131-164.
- Merton, Robert, 1974, On the pricing of corporate debt: the risk structure of interest rates, *Journal of Finance* 29, 449–470.
- Mitton, T. and K. Vorkink (2007), Equilibrium Underdiversification and the Preference for Skewness, *Review of Financial Studies* 20, 1255:1288.
- Nagel, Stefan, 2005, Short sales, institutional investors and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309.
- Ni, Sophie, 2009, Stock Option Returns: A Puzzle, working paper, Hong Kong University of Science and Technology.
- Ni, Sophie and Michael Lemmon, 2009, The Effects of Sentiment on Option Speculative Trading and Prices of Stock and Index Options, working paper, University of Utah and Hong Kong University of Science and Technology.
- Ohlson, James A., 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of Accounting Research* 18, 109-131.
- Shumway, Tyler, 2001, Forecasting bankruptcy more accurately: a simple hazard model," *Journal of Business* 74, 101-124.
- Tversky, Amos and Daniel Kahneman, 1992, Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty*, 5, 297-323.

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

Appendix 1: Definition of key variables

| Name | Description |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Jackpot _t | 1 if firm has continuously compounded returns > 100% over months $t+1$ to $t+12$, 0 otherwise. For firms that delist within the next 12 months we use returns over as many months as are available, adjusted for any delisting returns on CRSP. |
| JACKPOTP | Predicted probability of jackpot return from out-of-sample regressions (specification as in Table 3, baseline model) |
| DEATHP | Predicted probability of distress from the main model in Campbell, Hilscher, Szilagyi (2008) (Table IV, 12 month lag, page 2913 of CHS). Note, this is in-sample and computed based on quarterly CompuStat data. |
| <i>The next set of variables are used to predict jackpots in our main specification</i> | |
| SKEW | Skewness of log daily returns over the last 3 months, centered around 0 |
| RET12 | Log return over the past year |
| SALEG | Sales growth in year 'y' is $\ln(\text{Sales}_y / \text{Sales}_{y-1})$ |
| AGE | Time (in years) since appearance on CRSP |
| TANG | Gross PPE/Total assets |
| TURN | Detrended stock turnover. Computed as in Chen, Hong, and Stein (2001), as average past 6-month turnover minus average past 18-month turnover |
| STDEV | Standard deviation of daily returns over the past 3 months, centered around zero |
| SIZE | Log (Market capitalization in thousands) |
| <i>The following variables are used to predict jackpots in Model 3 and are defined in a similar manner as CHS (except we use annual rather than quarterly COMPUSTAT data)</i> | |
| ROMA | Return on Market Assets = Net Profits / (Market Equity + Total Liabilities) |
| MLEV | (Short term + Long term Debt) / (Market equity + Short term + Long term Debt) |
| RELSIZE | Log (Firm market cap / market cap of S&P 500) |
| CASH | Cash / (Market cap + Book value of Total liabilities) |
| PRC15 | Log of stock price, truncated at \$15 |
| EXRAVG | Excess returns of the stock over the S&P 500 over the last 12 months with geometrically declining weights |
| MB | Market to Book. Book Value of Equity is computed as in CHS, based on the procedure in Cohen, Polk, and Vuolteenaho (2003) |
| Note: All accounting data is lagged by six months, to ensure that they are known by investors. | |

Appendix 2: The Accuracy Ratio

The accuracy ratio is used to assess the predictive ability of a model that ranks elements. It has been used extensively in the credit risk literature in assessing the performance of credit rating models (e.g. Vassalou and Xing(2004)).

Let us suppose that a model ranks firms according to a measure of predicted jackpot probability. There are N firms in total in our sample and M of those will realize a jackpot return in the next one year. Let $\theta = \frac{M}{N}$ be the percentage of firms that will realize a jackpot return. For every integer λ between zero and 100, K_λ is the number of firms that will realize a jackpot return within $\lambda\%$ of firms with the highest jackpot probability. Of course, K_λ cannot be more than M . $f(\lambda)$ is defined as K_λ/M . Then $f(\lambda)$ takes values between zero and one, and is an increasing function of λ . Moreover, $f(0)=0$ and $f(100)=1$.

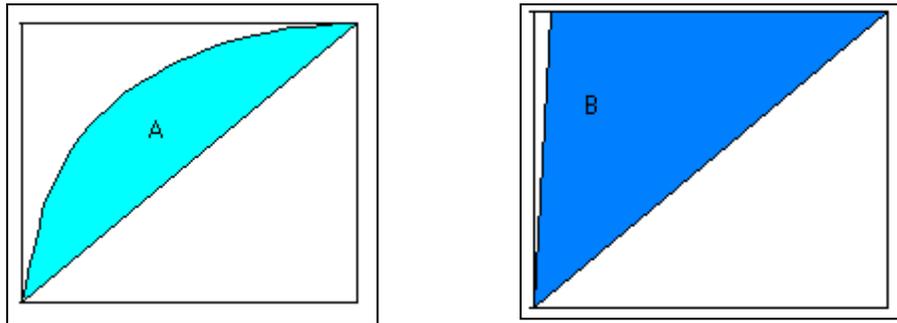
Suppose we had the “perfect measure” of future jackpot probability, and we were ranking stocks according to that. Our model would then have been able to perfectly predict jackpot for each integer λ , and $f(\lambda)$ would be given by:

$$f(\lambda) = \frac{\lambda}{\theta} \quad \text{for } \lambda < \theta \text{ and,}$$
$$f(\lambda) = 1 \quad \text{for } \lambda \geq \theta$$

The graph of average $f(\lambda)$ over all months in the sample for this perfect measure is shown as the kinked line in the right hand side panel of Figure A.1. At the other extreme, suppose we had zero information about future jackpot probability, and we were ranking the stocks randomly. If we did that a large number of times, $f(\lambda)$ would be equal to λ . Graphically, the average $f(\lambda)$ would correspond to the 45 degree line in the graphs of Figure A.1.

We measure the amount of information in a model by how far the graph of the average $f(\lambda)$ for a given model lies above the 45 degree line. Specifically, we measure it by the area between the 45-degree line and the graph of average $f(\lambda)$. This is depicted as the area A in Figure A.1. The accuracy ratio of a model is then defined as the ratio between the area associated with that model's average $f(\lambda)$ function, and the one associated with the "perfect" model's average $f(\lambda)$ function (area B in Figure A.1). Under this definition, the "perfect" model has accuracy ratio of one, and the zero-information model has an accuracy ratio of zero.

Figure A.1: Illustrating the accuracy ratio



Appendix 3: Correlated independent variables in Fama-MacBeth regressions

We simulate the cross-section of returns to understand the impact of adding correlated variables into Fama-Macbeth specifications. We consider two correlated variables, the probability of default, DEATHP, and the probability of jackpot returns, JACKPOTP. Expected returns are related to JACKPOTP but not to DEATHP as described below. Both DEATHP and JACKPOTP are measured with error and these observed values are used as explanatory variables in Fama- Macbeth regressions. We present results for calibrations with different degrees of correlation between DEATHP and JACKPOTP and different degrees of measurement error.

The specification:

1. True DEATHP and JACKPOTP are bivariate normal with unit variance and correlation (or covariance) ρ

$$\begin{pmatrix} DEATHP_{i,t} \\ JACKPOTP_{i,t} \end{pmatrix} \sim N \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

2. These are both observed with error:

$$\widehat{DEATHP}_{i,t} = DEATHP_{i,t} + u_{i,t}$$

$$\widehat{JACKPOTP}_{i,t} = JACKPOTP_{i,t} + v_{i,t}$$

Where:

$u_{i,t} \sim N(0, \sigma_u^2)$, and $v_{i,t} \sim N(0, \sigma_v^2)$ are measurement errors in DEATHP and JACKPOTP, respectively.

3. The data generating process for returns is:

$$r_{i,t} = r_{m,t} + \lambda_t JACKPOTP_{i,t} + \varepsilon_{i,t}$$

$r_{m,t} \sim N(0.01, 0.057735)$, is the monthly return on the average stock, with standard deviation chosen to be 20% per year.

$\varepsilon_{i,t} \sim N(0, \sigma_{ivol,i,t}^2)$, is an idiosyncratic shock, whose log variance is distributed normally across firms, with parameters calibrated to data (Mean=- 4.17 and Variance=1.38)

$\lambda_t \sim N(\mu_\lambda, \sigma_\lambda^2)$, is a time-varying premium for JACKPOTP.

We use this data generating process to simulate a panel of returns with 2000 firms and T=400 months. For each panel we run Fama-Macbeth regressions of returns on observed DEATHP and observed JACKPOTP. We report statistics over 500 such simulations.

Although we have several free parameters, these are restricted by observed data.

1. We tune $\mu_\lambda, \sigma_\lambda^2$ such that we match mean coefficients and t-statistics observed in the data in Fama-MacBeth regressions of returns on JACKPOTP (approximately -0.5 for the mean, and -5 for the t-statistic).
2. Since we can observe r_m , we use the average measured variance of market returns. It is comforting that the t-stats for the intercept in our calibration are reasonably close to the data.
3. We have three free parameters left: ρ , the true correlation between DEATHP and JACKPOTP, and σ_u^2, σ_v^2 , the measurement errors in DEATHP and JACKPOTP. However, the measured correlation between DEATHP and JACKPOTP ties down one degree of freedom for these three. Therefore we present results for a range of true ρ and true σ_v^2 and back out the σ_u^2 that will give a measured correlation between DEATHP and JACKPOTP of 0.4. Also, given the lower R^2 's in predicting jackpot returns as compared to those in predicting default in the data, it is likely that the measurement error in JACKPOTP is greater than that in DEATHP.

Table A1 reports results of these simulations for σ_u taking on values 0.1, 0.3, 0.5, and 0.7 and ρ taking values 0.5, 0.6, 0.7 and 0.8. We report three specifications: returns on both DEATHP and JACKPOTP, on only DEATHP and only JACKPOTP. All specifications include a constant as well.

We can see from the table that there are parameter values for which a high correlation between DEATHP and JACKPOTP, coupled with measurement error, results in significant t-statistics for both variables in Fama-MacBeth regressions. In many specifications (particularly those with higher measurement error for JACKPOTP than for DEATHP) the magnitudes of the coefficients for both DEATHP and JACKPOTP are roughly similar. Thus, Fama-MacBeth regressions may not have power to distinguish between the effects of jackpot probability and default probability on expected returns.

Table 1: Motivation

This table reports results for portfolios formed from sorts on DEATHP, the probability of default from the model in Campbell, Hilsher and Szilagyi (2008). Portfolios are value-weighted, formed monthly and skip a month between portfolio formation and measuring returns. Panel A reports excess returns and four factor (Fama-French three factors and momentum) alphas for these portfolios, the fraction of firms in each portfolio that realize annual returns greater than three benchmarks over the next year, and the average skewness of daily returns of each stock over the next 12 months. The benchmarks are log returns over 100%, arithmetic returns over 100% and, arithmetic returns over 75% over the next year. Panel B calibrates the Barberis and Huang (2008) model. The payoffs of the lottery asset are L with probability q , 1 otherwise. We set $L=5$ and $l=1$. For a given q , we search for a price P , such that the heterogeneous holding equilibrium exists. We report P , L/p , l/P , and $E(r-rf)$, the expected excess returns of the lottery asset. The sample period is 1972-2009.

Panel A: Key statistics of DEATHP sorted portfolios

| Decile | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 1 - 10 |
|---------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| Excess Return | 0.62 | 0.53 | 0.55 | 0.45 | 0.41 | 0.41 | 0.35 | 0.24 | 0.00 | -0.67 | 1.28 |
| t-stat | 2.74 | 2.44 | 2.40 | 1.96 | 1.65 | 1.54 | 1.18 | 0.72 | -0.01 | -1.44 | 3.40 |
| 4-factor Alpha | 0.11 | 0.01 | 0.11 | 0.05 | -0.05 | 0.01 | 0.01 | -0.19 | -0.33 | -1.10 | 1.21 |
| t-stat | 1.20 | 0.09 | 1.37 | 0.68 | -0.61 | 0.12 | 0.08 | -1.15 | -1.84 | -4.46 | 4.32 |
| Log ret>100% | 1.4% | 1.4% | 1.5% | 1.5% | 1.6% | 1.6% | 1.8% | 2.0% | 2.6% | 3.9% | -2.5% |
| Arithmetic returns > 100% | 4.6% | 4.8% | 4.9% | 4.8% | 5.0% | 5.2% | 5.6% | 6.0% | 7.1% | 9.1% | -4.5% |
| Arithmetic returns > 75% | 8.1% | 8.1% | 8.2% | 8.1% | 8.4% | 8.8% | 9.3% | 9.5% | 10.7% | 12.5% | -4.5% |
| Skewness (t+1,t+12) | 0.68 | 0.58 | 0.62 | 0.67 | 0.75 | 0.84 | 0.94 | 1.03 | 1.25 | 1.58 | -0.90 |
| Median Inst. Own. | 42.7% | 45.2% | 42.7% | 39.6% | 35.6% | 30.6% | 24.5% | 18.7% | 14.5% | 12.5% | 30.2% |

Panel B: Barberis and Huang (2008) calibration

| q | P | L/P | l/P | $E(r-rf)$ |
|------|------|-------|-------|-----------|
| 0.1 | 1.38 | 3.63 | 0.73 | -0.33% |
| 0.09 | 1.35 | 3.70 | 0.74 | -1.28% |
| 0.08 | 1.32 | 3.78 | 0.76 | -2.22% |
| 0.07 | 1.29 | 3.86 | 0.77 | -3.14% |
| 0.06 | 1.27 | 3.95 | 0.79 | -4.00% |
| 0.05 | 1.23 | 4.05 | 0.81 | -4.78% |
| 0.04 | 1.20 | 4.16 | 0.83 | -5.43% |
| 0.03 | 1.16 | 4.29 | 0.86 | -5.86% |
| 0.02 | 1.12 | 4.45 | 0.89 | -5.90% |
| 0.01 | 1.07 | 4.66 | 0.93 | -5.10% |

Table 2: Summary statistics

This table reports summary statistics for key variables used in this paper. These variable are defined in Appendix 1. Panel A reports statistics are for all firm-months where data is available for all variables. Panel B reports statistics for the sub-sample of firms that will realize a jackpot return (log return >100%) over the next twelve months measured from June each year. The sample period is 1951-2009.

| Panel A: Summary statistics for key variables | | | | | | | | |
|------------------------------------------------------------|-----------|--------|--------|-------|-----------|---------|-------|--------|
| | SKEW | RET12 | AGE | TANG | SALESGRTH | TURN | STDEV | SIZE |
| MEAN | 0.213 | -0.007 | 15.773 | 0.532 | 0.170 | 0.00162 | 0.035 | 4.753 |
| STD | 1.006 | 0.574 | 14.671 | 0.348 | 0.292 | 0.082 | 0.027 | 2.062 |
| MIN | -7.734 | -8.007 | 0.496 | 0.048 | -0.268 | -5.156 | 0.001 | -4.550 |
| MAX | 7.809 | 4.416 | 83.915 | 1.216 | 0.983 | 14.829 | 0.716 | 13.309 |
| N | 1,824,924 | | | | | | | |
| Panel B: Summary statistics for 'Jackpot' Subsample | | | | | | | | |
| | SKEW | RET12 | AGE | TANG | SALESGRTH | TURN | STDEV | SIZE |
| MEAN | 0.273 | -0.133 | 9.764 | 0.462 | 0.197 | 0.00174 | 0.051 | 3.584 |
| STD | 1.011 | 0.721 | 9.132 | 0.323 | 0.354 | 0.079 | 0.033 | 1.666 |
| MIN | -6.988 | -3.928 | 0.496 | 0.048 | -0.268 | -1.276 | 0.006 | -0.717 |
| MAX | 5.564 | 2.627 | 78.918 | 1.216 | 0.983 | 1.982 | 0.564 | 10.927 |
| N | 3136 | | | | | | | |

Table 3: In-sample predictors of jackpot returns

This table reports annual in-sample logit regressions of 'jackpot', a dummy variable that equals one if a stock's log return over the next 12 months (July to June) exceeds 100%, on a set of predictive variables from 1951-2009. The key predictive variables are defined in Appendix 1. All predictive variables are known as of end June, with accounting data lagged by six months to ensure availability. In addition, NASDAQ, SMALLDUMMY, and MEDIANDUMMY are dummy variables that take the value of 1 for Nasdaq listed firms, firms in the bottom and median tercile of market capitalization, respectively.

| Panel A: Baseline Model | | | | | |
|------------------------------------|-----------------|--------------------|----------------|--------------------------------------------------------------------------|----------------------|
| | Variable | Coefficient | T-stats | %change in odds ratio for a 1 σ change in X | R² |
| | Intercept | -3.29 | -36.00 | | 5.76% |
| | SKEW | 0.06 | 3.37 | 7.40% | |
| | RET12 | 0.18 | 4.42 | 9.10% | |
| | AGE | -0.02 | -8.75 | -27.60% | |
| | TANG | -0.25 | -3.45 | -8.20% | |
| | SALESGRTH | 0.29 | 4.10 | 8.20% | |
| | TURN | -0.43 | -2.18 | -3.50% | |
| | STDEV | 0.99 | 16.43 | 32.90% | |
| | SIZE | -0.22 | -15.94 | -34.90% | |
| Panel B: Alternative Models | | | | | |
| Model 2 | Intercept | -5.18 | -56.48 | | 6.07% |
| | SKEW | 0.06 | 3.17 | 7.60% | |
| | RET12 | 0.17 | 4.27 | 7.90% | |
| | AGE | -0.02 | -8.03 | -27.70% | |
| | TANG | -0.13 | -1.86 | -4.20% | |
| | SALESGRTH | 0.25 | 3.60 | 6.40% | |
| | TURN | -0.37 | -1.91 | -3.20% | |
| | STDEV | 0.88 | 13.81 | 29.20% | |
| | NASDAQ | 0.15 | 2.97 | 7.30% | |
| | SMALLDUMMY | 1.18 | 16.51 | 45.00% | |
| | MEDIANDUMMY | 0.77 | 10.61 | 40.70% | |
| Model 3 | Intercept | -5.60 | -21.23 | | 6.39% |
| | ROMA | 0.02 | 0.06 | 0.00% | |
| | MLEV | 0.10 | 0.93 | 2.00% | |
| | RELSIZE | -0.19 | -11.06 | -31.10% | |
| | CASH | 0.51 | 4.21 | 6.70% | |
| | EXRAVG | 3.98 | 10.39 | 21.40% | |
| | STDEV | 0.63 | 9.09 | 19.30% | |
| | MB | 0.05 | 7.05 | 15.40% | |
| | PRC15 | -0.50 | -9.91 | -24.10% | |

Table 4: Portfolios formed from a univariate sort on out-of-sample predicted jackpot probability

This table presents statistics of portfolios formed from decile sorts on predicted jackpot probability (JACKPOTP). JACKPOTP is from out-of-sample, expanding window, logit regressions of our baseline model (Table 2). Panel A reports excess returns and alphas of these portfolios from CAPM, Fama-French, and four factor (Fama-French and momentum) regressions. Panel B presents portfolio loadings in the four factor regression, while Panel C presents the characteristics of these portfolios. The sample period is 1972-2009.

| Panel A: 4 factor alphas (in % per month) of value-weighted portfolios sorted on JACKPOT | | | | | | | | | | | |
|-------------------------------------------------------------------------------------------------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------------|
| Decile | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 1 minus 10 |
| Excess Return | 0.44 | 0.52 | 0.54 | 0.61 | 0.53 | 0.56 | 0.45 | 0.24 | 0.03 | -0.62 | 1.06 |
| t-stat | 2.29 | 2.15 | 1.97 | 2.04 | 1.61 | 1.53 | 1.14 | 0.58 | 0.06 | -1.29 | 2.63 |
| CAPM Alpha | 0.07 | 0.05 | 0.02 | 0.06 | -0.06 | -0.08 | -0.22 | -0.44 | -0.69 | -1.32 | 1.39 |
| t-stat | 1.31 | 0.84 | 0.21 | 0.48 | -0.43 | -0.44 | -0.99 | -1.84 | -2.54 | -4.11 | 3.84 |
| 3-factor Alpha | 0.09 | 0.07 | 0.03 | 0.05 | -0.04 | -0.01 | -0.16 | -0.40 | -0.64 | -1.30 | 1.39 |
| t-stat | 2.43 | 1.10 | 0.39 | 0.50 | -0.41 | -0.11 | -1.28 | -2.61 | -3.64 | -6.22 | 6.28 |
| 4-factor Alpha | 0.04 | 0.11 | 0.12 | 0.06 | 0.01 | 0.08 | -0.06 | -0.24 | -0.45 | -1.06 | 1.10 |
| t-stat | 0.98 | 1.80 | 1.67 | 0.65 | 0.16 | 0.82 | -0.44 | -1.37 | -2.48 | -5.01 | 4.85 |
| Panel B: Factor Loadings in the four-factor Model | | | | | | | | | | | |
| MKT | 0.91 | 1.03 | 1.10 | 1.14 | 1.18 | 1.20 | 1.24 | 1.23 | 1.26 | 1.17 | -0.26 |
| t-stat | 70.13 | 60.43 | 49.85 | 40.94 | 52.37 | 42.35 | 32.38 | 25.18 | 23.60 | 16.94 | -3.54 |
| SMB | -0.26 | 0.07 | 0.30 | 0.56 | 0.76 | 0.90 | 1.11 | 1.18 | 1.39 | 1.62 | -1.88 |
| t-stat | -18.40 | 2.83 | 6.90 | 13.61 | 21.84 | 19.50 | 23.42 | 18.68 | 12.68 | 13.20 | -14.70 |
| HML | 0.04 | -0.06 | -0.12 | -0.09 | -0.23 | -0.35 | -0.37 | -0.39 | -0.45 | -0.48 | 0.52 |
| t-stat | 1.30 | -1.48 | -3.50 | -1.98 | -5.76 | -6.59 | -6.08 | -4.14 | -4.13 | -3.13 | 3.01 |
| WML | 0.05 | -0.04 | -0.09 | -0.01 | -0.05 | -0.10 | -0.10 | -0.16 | -0.18 | -0.23 | 0.28 |
| t-stat | 2.76 | -2.19 | -4.74 | -0.43 | -1.82 | -2.75 | -1.78 | -2.75 | -2.51 | -2.95 | 3.20 |
| Panel C: Portfolio Characteristics | | | | | | | | | | | |
| Portfolio STD | 4.79 | 4.65 | 4.91 | 4.86 | 5.26 | 5.74 | 6.36 | 6.96 | 8.28 | 9.86 | 7.47 |
| Portfolio Skew | -0.18 | -0.15 | -0.25 | -0.25 | -0.44 | -0.41 | -0.30 | -0.30 | 0.13 | 0.70 | -0.48 |

Table 5: The jackpot effect and limits to arbitrage

This table presents evidence on the effects of limits to arbitrage on the low average returns of jackpot stocks. We consider three measures of limits to arbitrage, size, price, institutional ownership, and analyst coverage. We first classify stocks into two groups, based on the level of the limits to arbitrage variable, and then examine returns of the portfolio long the highest jackpot probability quintile and short the lowest jackpot probability quintile within each group. Both long and short portfolios are value-weighted. We present mean returns of this portfolio, CAPM and four factor alphas, and mean relative size of each group. We also present the difference in mean JACKPOTP between the lowest and highest quintile of JACKPOTP within each group. Small (Big) stocks are smaller (larger) than the 30th (70th) NYSE size percentile. For Institutional Ownership and Analyst Coverage, we first measure the residual of each variable in a regression on RELSIZE and time dummies. 'Low' Institutional Ownership or Analyst Coverage stocks are in the smallest tercile, while 'High' stocks are in the largest tercile of their respective residual. Size results are for 1972-2009, while Institutional Ownership and Analyst coverage are for 1980-2009. All portfolios are value-weighted and skip a month between portfolio formation and measuring returns.

| Decile | Size | | Inst. Ownership | | Analyst Coverage | |
|----------------------|--------|-------|-----------------|-------|------------------|-------|
| | Small | Big | Low | High | Low | High |
| Mean Returns | -1.18 | -0.05 | -1.23 | -0.33 | -0.42 | -0.37 |
| t-stat | -3.95 | -0.19 | -2.76 | -0.86 | -0.90 | -0.85 |
| CAPM Alpha | -1.44 | -0.37 | -1.70 | -0.67 | -0.86 | -0.82 |
| t-stat | -5.32 | -1.48 | -4.21 | -1.89 | -2.59 | -2.09 |
| 4-factor Alpha | -0.86 | 0.01 | -1.19 | -0.41 | -0.54 | -0.25 |
| t-stat | -2.96 | 0.05 | -4.45 | -1.59 | -1.52 | -0.79 |
| Mean JACKPOTP spread | 3.8% | 0.6% | 2.4% | 3.1% | 2.2% | 2.4% |
| Mean RELSIZE | -11.01 | -6.83 | -9.58 | -9.61 | -9.76 | -9.22 |

Table 6: Firm Characteristics

This table presents average individual firm characteristics for portfolios sorted on predicted default probability according to the Campbell, Hilsher and Szilagyi (2008) model in Panel A and for portfolios sorted on out-of-sample predicted jackpot probability in Panel B. The variables are defined as in Appendix 1. Realized jackpot is the average of the binary variable jackpot, that is 1 if log returns over the next 12 months are greater than 100%. The sample period is 1972-2009.

| Panel A: Individual stock characteristics of DEATHP sorted portfolio | | | | | | | | | | |
|-------------------------------------------------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| JACKPOTP | 0.9% | 0.9% | 0.9% | 1.0% | 1.2% | 1.3% | 1.5% | 1.8% | 2.2% | 3.4% |
| Realized jackpot | 1.4% | 1.4% | 1.5% | 1.5% | 1.6% | 1.6% | 1.8% | 2.0% | 2.6% | 3.9% |
| RET12 | 26.0% | 21.6% | 17.2% | 13.2% | 9.1% | 4.7% | 0.2% | -6.0% | -18.8% | -44.9% |
| SIZE | 5.42 | 5.75 | 5.58 | 5.36 | 5.12 | 4.84 | 4.46 | 4.06 | 3.57 | 2.93 |
| BM | 0.62 | 0.59 | 0.66 | 0.72 | 0.79 | 0.85 | 0.92 | 0.99 | 1.11 | 1.27 |
| SALESGRTH | 14.2% | 17.1% | 17.9% | 17.9% | 18.0% | 17.7% | 17.6% | 17.2% | 16.7% | 14.5% |
| MLEV | 5.9% | 11.0% | 15.1% | 18.6% | 21.6% | 24.4% | 27.1% | 30.1% | 33.9% | 39.4% |
| SKEW | 0.32 | 0.24 | 0.22 | 0.21 | 0.20 | 0.19 | 0.20 | 0.20 | 0.18 | 0.16 |
| Panel B: Individual stock characteristics of JACKPOTP sorted portfolio | | | | | | | | | | |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| JACKPOTP | 0.2% | 0.4% | 0.5% | 0.7% | 0.9% | 1.1% | 1.4% | 1.8% | 2.4% | 4.6% |
| Realized jackpot | 0.1% | 0.3% | 0.5% | 0.8% | 1.3% | 1.8% | 2.2% | 2.7% | 3.1% | 3.6% |
| RET12 | 9.6% | 9.9% | 10.2% | 10.2% | 9.3% | 7.4% | 5.5% | 2.5% | -3.6% | -16.5% |
| SIZE | 8.10 | 6.89 | 6.19 | 5.62 | 5.14 | 4.73 | 4.35 | 3.98 | 3.56 | 2.94 |
| BM | 0.74 | 0.74 | 0.75 | 0.77 | 0.79 | 0.81 | 0.83 | 0.85 | 0.89 | 1.00 |
| SALESGRTH | 9.1% | 11.1% | 13.0% | 14.5% | 16.2% | 17.7% | 19.2% | 21.1% | 23.9% | 27.1% |
| MLEV | 25.9% | 24.5% | 23.1% | 22.4% | 22.2% | 22.1% | 21.8% | 21.2% | 20.9% | 22.2% |
| SKEW | -0.01 | 0.04 | 0.08 | 0.13 | 0.17 | 0.20 | 0.24 | 0.28 | 0.33 | 0.43 |

Table 7: The relation between returns of distress and jackpot strategies

Panel A presents Spearman correlations between predicted jackpot probability (JACKPOTP) and predicted default probability (DEATHP). JACKPOTP is from the different models in Table 3, with different predictor variables (models 2-4) and different cut-offs used in defining jackpot (arithmetic returns of 50%, 75% and 100% over the subsequent year). Panel B presents time-series regressions of returns of the 'distress strategy' on different factors. The distress strategy, DEATHLS, is long the bottom 10% of stocks (lowest default probability) and short the top 10% of stocks (highest default probability) according to the default probability model in Campbell, Hilsher and Szilagyi (2008). Specification 1 regresses the distress strategy on JACKPOTLS. JACKPOTLS is a portfolio of stocks that is long the decile least likely to achieve jackpot returns and short the decile most likely to achieve jackpot returns, according to the baseline out-of-sample jackpot prediction model. Specification 2 uses both the four factor model as well as returns to the jackpot strategy as explanatory variables. Panel C regresses JACKPOTLS on different factors. The first specification regress JACKPOTLS on DEATHLS, and the next specification adds in the four factors. The sample period is 1972-2009. All portfolios are value-weighted and skip a month between portfolio formation and measuring returns.

Panel A: Spearman correlations with predicted default probability from CHS (DEATHP)

| | <i>Correlation</i> |
|------------------|--------------------|
| JACKPOTP | 41.80% |
| JACKPOTP Model 2 | 39.54% |
| JACKPOTP Model 3 | 39.26% |
| JACKPOTP50 | 45.43% |
| JACKPOTP75 | 45.26% |
| JACKPOTP100 | 45.64% |

Panel B: Explaining the returns of the distress strategy

| | <i>Coef.</i> | <i>t-value</i> | <i>Coef.</i> | <i>t-value</i> |
|-----------|--------------|----------------|--------------|----------------|
| Intercept | 0.54 | 1.47 | 0.71 | 2.69 |
| MKT | | | -0.35 | -5.06 |
| SMB | | | -0.57 | -3.45 |
| HML | | | -0.38 | -2.98 |
| WML | | | 0.79 | 6.68 |
| JACKPOTLS | 0.56 | 5.18 | 0.34 | 5.37 |
| R-square | 32.54% | | | 66.18% |

Panel C: Explaining the returns of the jackpot strategy

| | <i>Coef.</i> | <i>t-value</i> | <i>Coef.</i> | <i>t-value</i> |
|-----------|--------------|----------------|--------------|----------------|
| Intercept | 0.40 | 1.19 | 0.81 | 2.86 |
| MKT | | | -0.03 | -0.40 |
| SMB | | | -1.14 | -7.00 |
| HML | | | 0.60 | 3.96 |
| WML | | | -0.07 | -0.73 |
| DEATHPLS | 0.58 | 10.01 | 0.36 | 4.66 |
| R-square | 32.54% | | | 65.21% |

Table 8: Distinguishing between the effects of distress and jackpot on expected returns

This table presents results for firms in the top quintile portfolio of DEATHP. We sort firms in the top predicted distress quintile based on their skewness in daily log returns over the past three months into two portfolios, high (greater than than the 70th percentile) and low (smaller than the 30th percentile). We also sort firms into 'traditional distress' (the bottom 30% of all firms in terms of sales growth and bottom 30% in M/B or with negative book equity) and 'speculative' (top 30% of all firms in terms of sales growth and top 30% of M/B) portfolios. We then compute value-weighted returns and four factor alphas for each portfolio, skipping a month between portfolio formation and measuring returns. We also report annualized mean DEATHP and mean JACKPOTP for each portfolio

| | Panel A: Skewness | | | Panel B: Sales Growth and M/B | | |
|----------------|--------------------------|--------|----------|--------------------------------------|--------|-------------|
| | Low | High | Low-High | Distress | Spec. | Dist.-Spec. |
| Mean return | 0.53% | 0.06% | 0.47% | 0.87% | -0.9% | 1.77% |
| 4 factor alpha | -0.21% | -0.75% | 0.54% | -0.14% | -1.70% | 1.55% |
| t-value | -1.18 | -3.3 | 2.21 | -0.55 | -4.81 | -3.68 |
| | 0.002 | 0.0022 | | 0.0023 | 0.002 | |
| Mean DEATHP | 2.28% | 2.40% | | 2.64% | 2.28% | |
| Mean JACKPOTP | 1.20% | 1.48% | | 1.09% | 1.97% | |
| N | 277 | 223 | | 169 | 53 | |

Table 9: Fama-MacBeth regressions

This table presents results of monthly Fama-MacBeth regressions. Explanatory variables include book-to-market (BM), size (SIZE), momentum returns from t-12 to t-2 (Ret (t-12,t-2)), stock return volatility (STDEV), out of sample predicted probability of jackpot (JACKPOTP) and predicted probability of distress (DEATHP). All variables (except returns) are winsorized at their 5th and 95th percentiles. All explanatory variables are standardized in each cross section, to have mean zero and standard deviation 1. Therefore, the coefficients can be interpreted as the effect of a 1 standard deviation shock of the explanatory variable on expected returns. All explanatory variables are known as of month 't-2', while returns on the left hand side are of month 't'. The sample is restricted to firm-months where JACKPOTP and DEATHP are both available and the time period is 1972-2009.

| Variable | (1) | | (2) | | (3) | | (4) | |
|----------------|--------|---------|--------|---------|--------|---------|--------|---------|
| | Mean | t-value | Mean | t-value | Mean | t-value | Mean | t-value |
| Intercept | 1.14% | 3.93 | 1.12% | 3.84 | 1.24% | 4.16 | 1.19% | 3.97 |
| LogBM | 0.29% | 5.08 | 0.31% | 5.64 | 0.26% | 4.71 | 0.28% | 5.24 |
| Ret (t-12,t-2) | 0.53% | 5.82 | 0.38% | 4.35 | 0.52% | 5.78 | 0.37% | 4.41 |
| LogSigma | -0.26% | -2.31 | -0.10% | -0.84 | -0.03% | -0.25 | 0.11% | 0.97 |
| Size | -0.18% | -2.42 | -0.23% | -3.16 | -0.38% | -5.24 | -0.40% | -5.47 |
| DEATHP | | | -0.42% | -8.09 | | | -0.41% | -7.63 |
| JACKPOTP | | | | | -0.37% | -4.96 | -0.33% | -4.27 |

Table 10: Alternate specifications and sample periods for portfolios formed from sorts on predicted JACKPOT

This table presents four-factor alphas of portfolios formed from alternate specifications (Panel A), time periods (Panel B) and, different definitions of jackpots (Panel C). Panel A reports four-factor alphas of portfolios formed from out-of-sample predicted jackpot probability based on the two alternate models with different predictor variables defined in Table 2. Panel B examines alphas from our baseline model over different sub-samples. Panel C reports alphas of portfolios formed based on out-of-sample predictions of jackpot probability with our baseline variables, with different definitions of jackpots (>50%, 75%, and 100% arithmetic returns over the next year). The sample period is 1972-2009.

| Panel A: Different models to forecast JACKPOT | | | | | | | | | | | |
|------------------------------------------------------|------|-------|------|-------|-------|-------|-------|-------|-------|-------|------------|
| Decile | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 1 minus 10 |
| Model 2 | 0.02 | -0.02 | 0.12 | 0.24 | -0.02 | 0.45 | -0.13 | -0.19 | -0.57 | -0.85 | 0.87 |
| t-stat | 0.23 | -0.27 | 1.51 | 2.07 | -0.15 | 2.96 | -0.96 | -1.16 | -3.26 | -4.25 | 3.58 |
| Model 3 | 0.07 | -0.02 | 0.10 | -0.14 | -0.01 | 0.10 | -0.02 | -0.14 | -0.05 | -0.62 | 0.68 |
| t-stat | 1.69 | -0.39 | 1.51 | -1.74 | -0.07 | 0.75 | -0.16 | -0.88 | -0.27 | -2.76 | 2.83 |
| Panel B: Sub-samples | | | | | | | | | | | |
| Decile | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 1 minus 10 |
| 1972-1979 | 0.01 | 0.14 | 0.21 | 0.01 | 0.12 | 0.07 | -0.15 | -0.35 | -0.32 | -1.11 | 1.12 |
| t-stat | 0.09 | 1.38 | 1.57 | 0.06 | 0.70 | 0.40 | -0.64 | -1.53 | -1.05 | -4.60 | 4.38 |
| 1980-1989 | 0.06 | 0.10 | 0.10 | 0.10 | 0.07 | 0.05 | -0.23 | -0.30 | -1.13 | -1.61 | 1.67 |
| t-stat | 1.16 | 1.42 | 1.05 | 0.77 | 0.72 | 0.31 | -1.79 | -2.01 | -6.72 | -5.71 | 5.60 |
| 1990-1999 | 0.06 | 0.05 | 0.08 | 0.02 | 0.01 | 0.01 | -0.25 | -0.06 | 0.06 | -0.83 | 0.89 |
| t-stat | 1.12 | 0.53 | 0.60 | 0.11 | 0.08 | 0.06 | -1.17 | -0.21 | 0.15 | -1.82 | 1.82 |
| 2000-2009 | 0.09 | 0.13 | 0.08 | 0.17 | -0.14 | 0.06 | 0.15 | -0.45 | -0.65 | -1.14 | 1.23 |
| t-stat | 0.97 | 0.93 | 0.46 | 0.95 | -0.78 | 0.23 | 0.50 | -1.38 | -1.87 | -2.52 | 2.64 |
| Panel C: Different definitions of JACKPOT | | | | | | | | | | | |
| Decile | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 1 minus 10 |
| JACKPOT as 50% | 0.05 | 0.15 | 0.13 | 0.05 | 0.03 | -0.08 | -0.19 | -0.39 | -0.67 | -1.07 | 1.12 |
| t-stat | 1.45 | 2.65 | 1.63 | 0.53 | 0.33 | -0.76 | -1.82 | -3.31 | -4.42 | -5.02 | 5.05 |
| JACKPOT as 75% | 0.04 | 0.17 | 0.05 | 0.05 | 0.12 | -0.06 | -0.17 | -0.30 | -0.69 | -1.03 | 1.08 |
| t-stat | 1.27 | 2.55 | 0.73 | 0.65 | 1.12 | -0.55 | -1.38 | -1.93 | -4.61 | -4.68 | 4.63 |
| JACKPOT as 100% | 0.05 | 0.10 | 0.18 | 0.00 | 0.01 | 0.09 | -0.12 | -0.36 | -0.62 | -1.02 | 1.07 |
| t-stat | 1.33 | 1.65 | 2.71 | 0.02 | 0.11 | 0.72 | -0.91 | -2.17 | -3.60 | -4.45 | 4.40 |

Figure 1: The time-series of mean jackpot probability

This figure plots mean realized jackpot probability for all stocks from 1951 to 2008. Jackpot is a binary variable equal to 1 if log returns over the next year for a stock are greater than 100%. For example, the point at 31 Dec 2008 reflects the fraction of stocks that have log returns greater than 100% from 01 Jan 2009 to 31 Dec 2009.

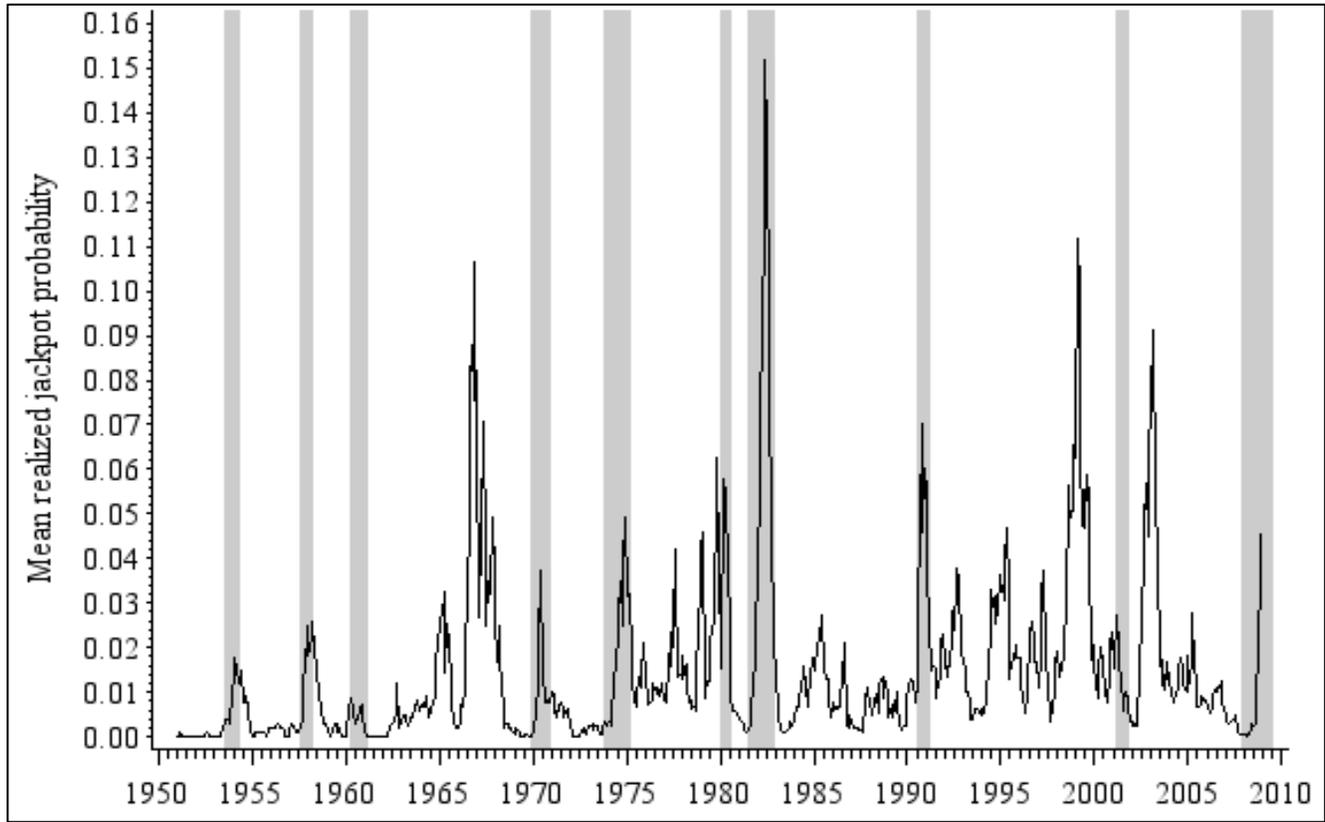
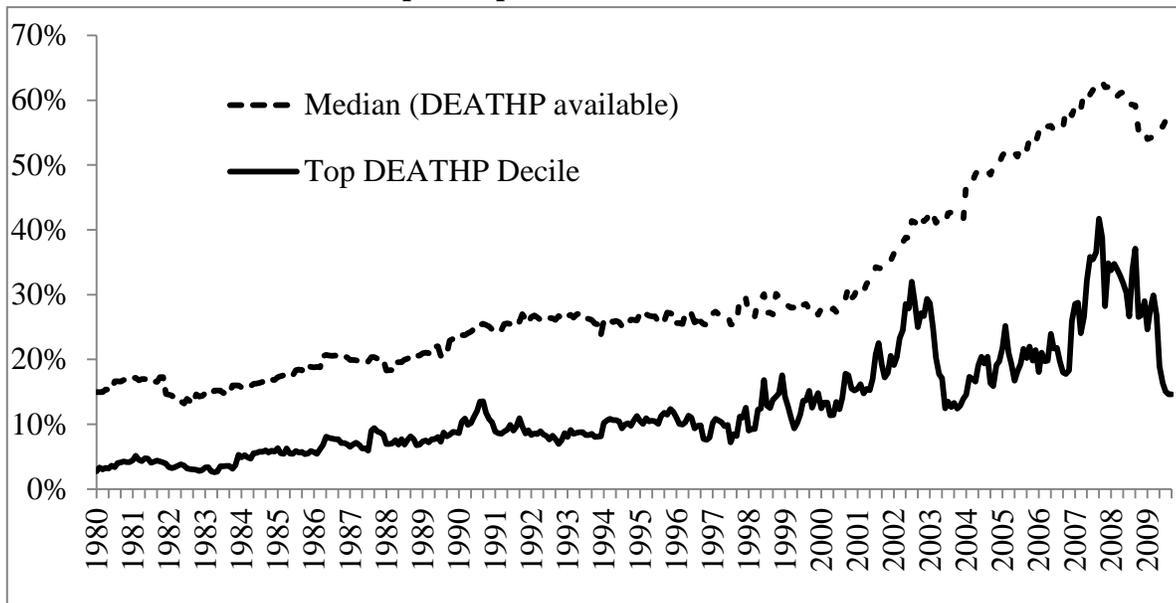


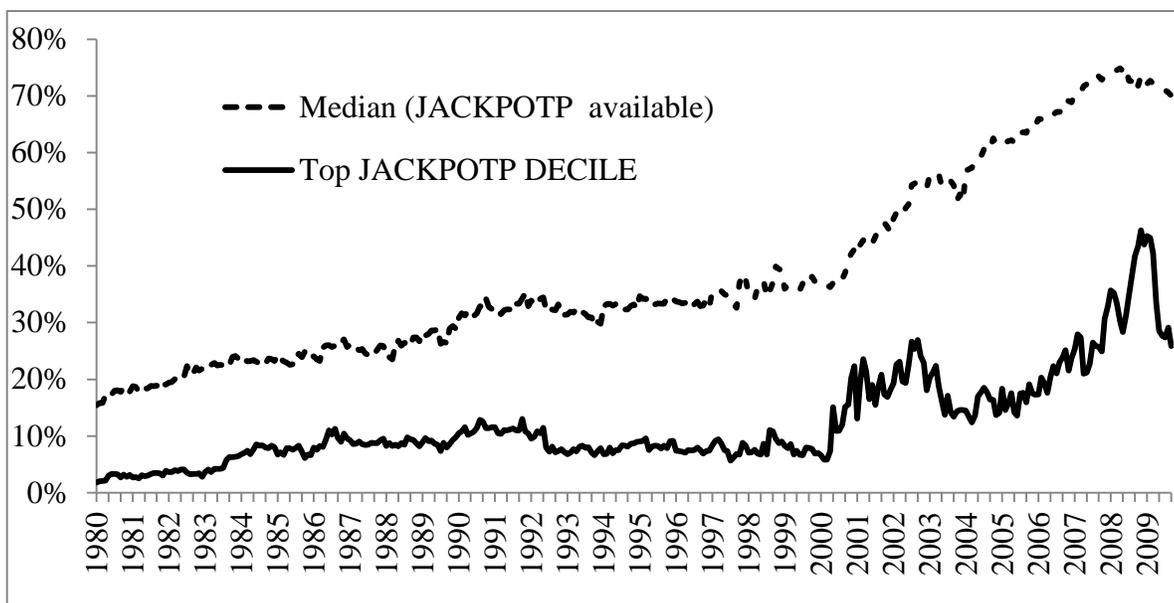
Figure 2: Institutional ownership, default, and jackpot probability

This figure plots institutional ownership for the top JACKPOTP and DEATHP decile portfolios. Panel A reports median institutional ownership for the top DEATHP decile along with median institutional ownership for all firms with sufficient data available to compute DEATHP. Panel B reports median institutional ownership for the top JACKPOTP decile, along with median institutional ownership for all firms with sufficient data available to compute JACKPOTP. Panels C and D report results for an event -study, where the event is the entry of a firm into the top DEATHP (Panel C) or JACKPOTP (Panel D). We present average excess institutional ownership of the firm over the mean institutional ownership that quarter, for four quarters before and four quarters after, entry into the top decile. In order to ensure a comparable sample, we restrict the sample to those firms that have data for the four prior quarters.

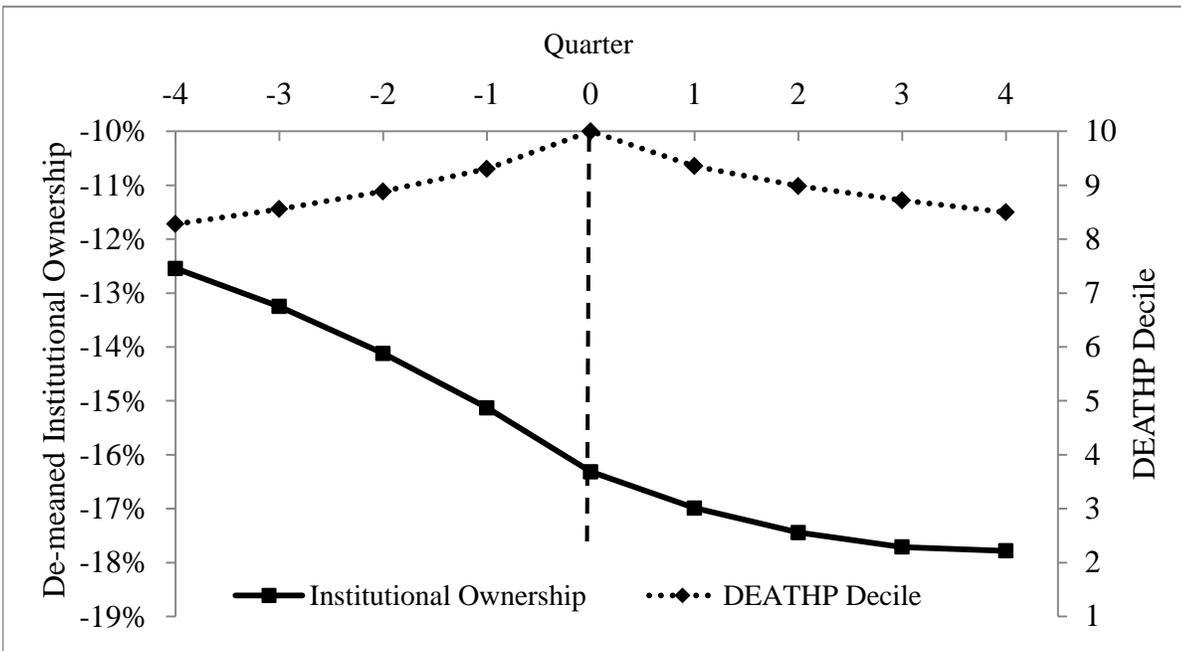
Panel A: Institutional Ownership for top DEATHP decile



Panel B: Institutional Ownership for top JACKPOTP decile



Panel C: Institutional ownership around entry into the top DEATHP decile



Panel D: Institutional ownership around entry into the top JACKPOTP decile

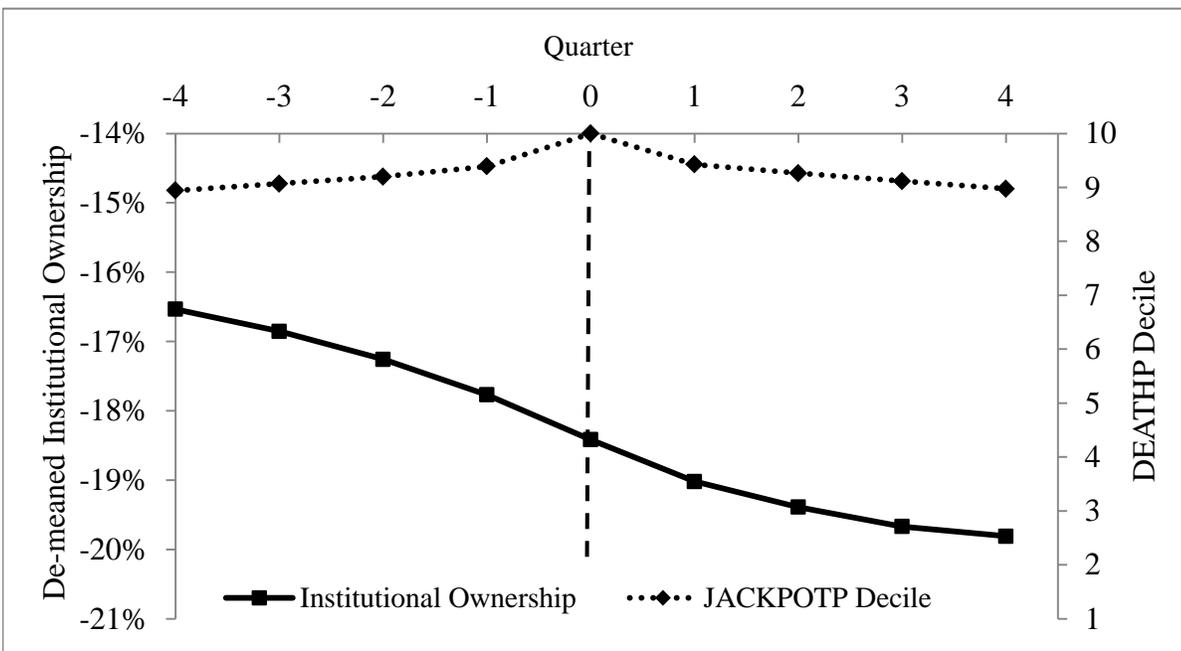


Figure 3: Carhart four-factor regression loadings for distress and jackpot strategies

This figure plots loadings on the factors in the Carhart four-factor model for both the ten DEATHP portfolios and the ten JACKPOTP portfolios. The estimation period is from 1972 to 2009.

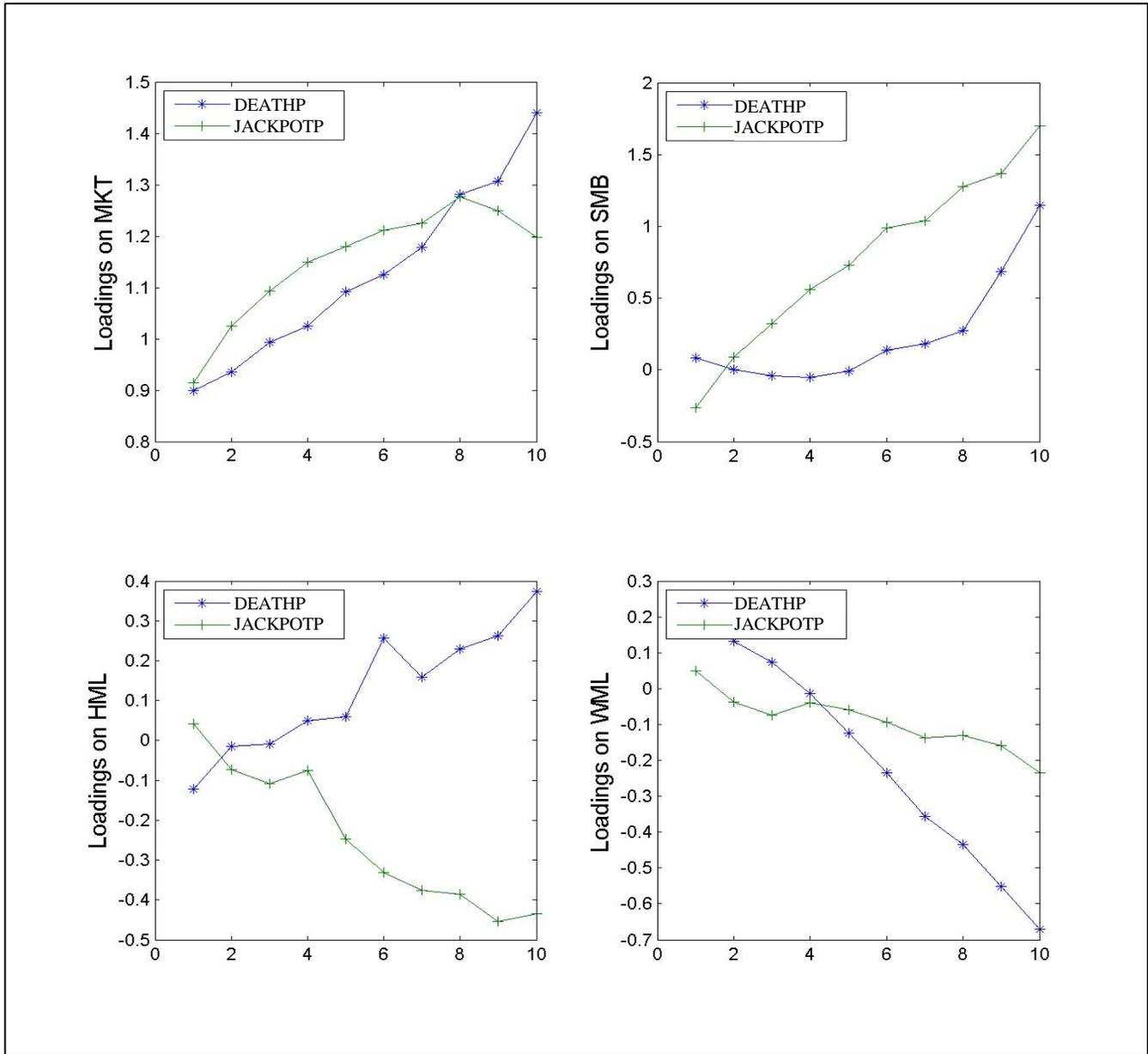


Table A1: The impact of correlated variables on Fama-Macbeth regressions

This table presents simulations of Fama-Macbeth regressions with correlated independent variables measured with error. The simulations are described in Appendix 3 and are run for different values for measurement error in death ($\sigma_u = 0.1, 0.3, 0.5,$ and 0.7) and the true correlation between JACKPOTP and DEATH ($\rho = 0.5, 0.6, 0.7,$ and 0.8). Panel A reports average coefficients across 500 simulations for a regression with both observed JACKPOT (Jackpothat) and observed DEATHP (Deathhat). Panel B has only Jackpothat while Panel C has only Deathhat. All three Panels include an intercept. Panel D reports measurement error in JACKPOTP that is used in each simulation. This is backed out from the measurement error in DEATHP and true correlation so that the observed correlation between Deathhat and Jackpothat is 0.4.

| | Intercept Jackpothat Deathhat | | | Intercept Jackpothat Deathhat | | | Intercept Jackpothat Deathhat | | | Intercept Jackpothat Deathhat | | |
|-----------------------------------------------------------------------|-------------------------------|--------|--------|-------------------------------|--------|--------|-------------------------------|--------|--------|-------------------------------|--------|--------|
| | σ_u | 0.1 | | 0.3 | | | 0.5 | | | 0.7 | | |
| True ρ | | | | | | | | | | | | |
| Panel A: Specification with intercept, Jackpothat and deathhat | | | | | | | | | | | | |
| 0.5 Coeff | 0.99% | -0.57% | -0.17% | 1.00% | -0.61% | -0.14% | 1.00% | -0.68% | -0.09% | 1.00% | -0.77% | -0.02% |
| t-stat | 3.82 | -5.44 | -5.03 | 3.86 | -5.48 | -4.84 | 3.86 | -5.56 | -4.21 | 3.85 | -5.50 | -1.27 |
| 0.6 Coeff | 1.00% | -0.40% | -0.31% | 1.00% | -0.44% | -0.28% | 1.02% | -0.51% | -0.23% | 1.00% | -0.58% | 0.16% |
| t-stat | 3.85 | -5.36 | -5.31 | 3.86 | -5.42 | -5.35 | 3.92 | -5.50 | -5.29 | 3.86 | -5.41 | -4.93 |
| 0.7 Coeff | 1.00% | -0.28% | -0.45% | 1.00% | -0.31% | -0.41% | 1.00% | -0.37% | -0.35% | 1.02% | -0.45% | -0.28% |
| t-stat | 3.85 | -5.34 | -5.48 | 3.84 | -5.42 | -5.46 | 3.83 | -5.37 | -5.35 | 3.92 | -5.48 | -5.37 |
| 0.8 Coeff | 1.02% | -0.18% | -0.57% | 1.03% | -0.20% | -0.53% | 0.98% | -0.26% | -0.46% | 1.00% | -0.33% | -0.39% |
| t-stat | 3.92 | -5.16 | -5.53 | 3.95 | -5.14 | -5.45 | 3.80 | -5.29 | -5.41 | 3.84 | -5.43 | -5.45 |
| Panel B: Specification with intercept and Jackpothat | | | | | | | | | | | | |
| 0.5 Coeff | 0.99% | -0.63% | | 1.00% | -0.66% | | 1.00% | -0.72% | | 1.00% | -0.78% | |
| t-stat | 3.82 | -5.46 | | 3.85 | -5.49 | | 3.86 | -5.57 | | 3.85 | -5.50 | |
| 0.6 Coeff | 1.00% | -0.53% | | 1.00% | -0.55% | | 1.02% | -0.60% | | 1.00% | -0.64% | |
| t-stat | 3.85 | -5.41 | | 3.86 | -5.46 | | 3.92 | -5.52 | | 3.85 | -5.43 | |
| 0.7 Coeff | 1.00% | -0.46% | | 1.00% | -0.48% | | 1.00% | -0.50% | | 1.02% | -0.56% | |
| t-stat | 3.85 | -5.48 | | 3.84 | -5.51 | | 3.83 | -5.43 | | 3.92 | -5.52 | |
| 0.8 Coeff | 1.02% | -0.41% | | 1.03% | -0.41% | | 0.98% | -0.44% | | 1.00% | -0.49% | |
| t-stat | 3.92 | -5.52 | | 3.95 | -5.42 | | 3.80 | -5.43 | | 3.84 | -5.51 | |

Panel C: Specification with intercept and deathhat

| | Intercept | Jackpothat | Deathhat |
|-------------------------------|-----------|------------|----------|-----------|------------|----------|-----------|------------|----------|-----------|------------|----------|
| σ_u | 0.1 | | | 0.3 | | | 0.5 | | | 0.7 | | |
| True ρ | | | | | | | | | | | | |
| 0.5 Coeff | 0.99% | | -0.39% | 1.00% | | -0.38% | 1.00% | | -0.36% | 0.99% | | -0.35% |
| t-stat | 3.82 | | -5.42 | 3.86 | | -5.43 | 3.86 | | -5.49 | 3.86 | | -5.40 |
| 0.6 Coeff | 1.00% | | -0.47% | 1.00% | | -0.46% | 1.02% | | -0.43% | 1.00% | | -0.39% |
| t-stat | 3.92 | | -5.49 | 3.86 | | -5.46 | 3.85 | | -5.40 | 3.85 | | -5.38 |
| 0.7 Coeff | 1.00% | | -0.56% | 1.00% | | -0.54% | 1.00% | | -0.49% | 1.02% | | -0.46% |
| t-stat | 3.83 | | -5.42 | 3.84 | | -5.51 | 3.85 | | -5.51 | 3.92 | | -5.50 |
| 0.8 Coeff | 1.02% | | -0.64% | 1.03% | | -0.61% | 0.98% | | -0.57% | 1.00% | | -0.52% |
| t-stat | 3.80 | | -5.45 | 3.95 | | -5.47 | 3.92 | | -5.55 | 3.84 | | -5.51 |

Panel D: Measurement error for Jackpot (σ_v)

| σ_u | 0.1 | | | 0.3 | | | 0.5 | | | 0.7 | | |
|-------------------------------|------|--|--|------|--|--|------|--|--|------|--|--|
| True ρ | | | | | | | | | | | | |
| 0.5 | 0.74 | | | 0.66 | | | 0.50 | | | 0.22 | | |
| 0.6 | 1.11 | | | 1.10 | | | 0.89 | | | 0.71 | | |
| 0.7 | 1.43 | | | 1.35 | | | 1.20 | | | 1.03 | | |
| 0.8 | 1.72 | | | 1.63 | | | 1.48 | | | 1.30 | | |