MARKET AND PROFESSIONAL DECISION-MAKING
UNDER RISK AND UNCERTAINTY
DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the
Graduate School of The Ohio State University

By
Erick Davidson, B.A., M.A.

*****

The Ohio State University
2007

Dissertation Committee: Approved by
Dr. Alan Randall, Adviser
Dr. Brian Roe
Dr. Mario Miranda
Dr. Abdoul Sam

-------------
Adviser
Graduate Program in
Agricultural, Environmental
and Development Economics
This dissertation explores decision making under risk and uncertainty by professionals and markets composed of professionals. Both essays use empirical data from some of the most competitive economic environments imaginable. The first essay looks at market prices resulting from the sum of both professional and lay choices while the second analyzes the individual choices of professional gamblers. Both essays propose a theoretical framework to not just positively identify what professionals do, but also prescribe normatively what they should do. In both cases, the two are found to be different.

The first essay, Market Response to Risk and Uncertainty, 2004 Hurricane Forecasts, develops a simple function to explain insurance losses from hurricanes as a function of short-term forecasts. After demonstrating the accuracy of the function in explaining 2004 insurance claims, the remainder of the essay looks at the stock market’s use of these forecasts in pricing insurer and US economy risk. Despite causing billions in damages, both hurricanes and hurricane forecasts are found to have only marginal impacts on financial markets. Surprisingly, markets fail to make efficient use of hurricane intensity in pricing both insurer and general market exposure to hurricane
risk. A potential explanation for market inefficiency around hurricane information is that, like researchers, financial actors may be confounding uncertainty for unpredictability.

The second essay, *Know When to Hold’em*, examines a specific decision within a highly popular, high-stakes version of poker. Like financial markets analyzed in the first essay, professional gamblers must make risky decisions with uncertain probabilities of success. Gamblers are found to be both overly conservative in their choices and overly confident in their abilities to predict uncertain outcomes. A simple statistical model that generalizes across situations to form a naïve probability of having the best cards, is found to be as effective in decision making as players’ true expectations of winning. Additionally, a dynamic theoretical model is presented in order to show professional divergence from risk-neutral expected profit maximization in the credit constrained world of tournament poker. Interestingly the value function, derived from this model, is equivalent to an optimal stock price of a poker player.
Dedicated to my entire family

with particular thanks to Sue and Meredith.
ACKNOWLEDGMENTS

This research was made possible by the Environmental Policy Initiative (EPI). Without either EPI’s financial support or the exposure to EPI’s interdisciplinary collaborations, I would not have been able to complete this work. Additionally I am grateful to Brian Roe for identifying high-stakes poker as a potentially rich domain of economic decision making. Mike Brady, Yun Jae Hwang, and Karl Muessen enabled research into televised poker through assistance in data coding. The poker analysis has also greatly benefited from conversations with Matt Roberts, Tim Haab, Audrey Light, Dan Levin, John Kagel, and Jim Peck. Finally and most importantly, I am thankful for the time and effort spent by the committee members; Alan Randall, Brian Roe, Mario Miranda, and Abdoul Sam. My dissertation has taken shape thanks to helpful revisions and numerous conversations with these men. Both the exposition and the substance of the analysis have been greatly improved and expanded by Alan’s and Brian’s critical analysis of both the stated and unstated hypotheses and Mario’s rigor in modeling dynamic problems.
VITA

August 9, 1974 ........................................... Born, Charleston, West Virginia, USA

1996 ........................................................ B.A. Political Science & French,
Central College

2004 ........................................................ M.A. Economics,
The Ohio State University

2003 - present ........................................... Graduate Research Assistant,
The Ohio State University

PUBLICATIONS


FIELDS OF STUDY

Major Field: Agricultural, Environmental, and Development Economics

Areas of Emphasis: Natural Resource Economics, Applied Econometrics, Risk Analysis
# TABLE OF CONTENTS

Abstract ........................................................................................................................................ ii  
Acknowledgments .................................................................................................................... v  
Vita ........................................................................................................................................ vi  
List of Tables ........................................................................................................................... viii  
List of Figures .......................................................................................................................... ix  

Chapter 1 ...................................................................................................................................... 1  
1.1 Abstract .............................................................................................................................. 1  
1.2 Introduction ........................................................................................................................ 2  
1.3 Data ..................................................................................................................................... 6  
1.4 2004 Hurricane Losses ...................................................................................................... 9  
1.5 2004 Stock Returns .......................................................................................................... 14  
1.6 Individual Insurer Returns .............................................................................................. 24  
1.7 Conclusions ...................................................................................................................... 28  
1.8 Caveats & Future Research ............................................................................................. 30  
References ...................................................................................................................................... 33  

Chapter 2 ..................................................................................................................................... 41  
2.1 Abstract .............................................................................................................................. 41  
2.2 Introduction ........................................................................................................................ 42  
2.3 Optimization in Poker ...................................................................................................... 47  
2.4 Empirics ............................................................................................................................ 61  
2.5 Outcomes ........................................................................................................................... 77  
2.6 Conclusion ........................................................................................................................ 88  
References ...................................................................................................................................... 90  

Appendix A: Poker Description/Game Tree .............................................................................. 101  
Appendix B: Poisson Trial ........................................................................................................ 103  

List of References .................................................................................................................... 104
LIST OF TABLES

Table 1.1 Regressions of forecasts on eventual losses ................................................... 11
Table 1.2 Realized and Predicted Hurricane Losses.......................................................... 13
Table 1.3 Regress Hurricane Info on Stock Returns.......................................................... 21

Table 2.1 Summary Empirical Estimation Results............................................................. 68
LIST OF FIGURES

Figure 1.1 Market Index Response to $\Delta$ forecast .................................................... 35
Figure 1.2 Insurance Index Response to $\Delta$ forecast ................................................... 36
Figure 1.3 Insurance Losses Response to $\Delta$ forecast ............................................... 37
Figure 1.4 Histogram of $\hat{\beta}_p$, impact of path forecasts on insurer returns .......... 38
Figure 1.5 Histogram of $\hat{\beta}_u$, impact of landfall on insurer returns ................... 39
Figure 1.6 Value of 2004 Insured Coastal Properties ...................................................... 40

Figure 2.1 Decision Tree – Poker final decision ............................................................. 51
Figure 2.2 Static indifference curves of loss averse and risk neutral poker player .... 93
Figure 2.3 Dynamic indifference curve of loss averse and risk neutral poker player ... 94
Figure 2.4 Graphical & Statistical Comparisons of Outcomes when player has best hand ................................................................. 95
Figure 2.5 Graphical & Statistical Comparisons of Outcomes when player has worst hand ........................................................................................................................................ 96
Figure 2.6 Graphical & Statistical Comparisons of Outcomes over all hands .......... 97
Figure 2.7 Value function of average cards with six players .......................................... 98
Figure 2.8 Value function of average cards with four players ......................................... 99
Figure 2.9 Value function of average cards with two players ........................................ 100
CHAPTER 1

MARKET RESPONSE TO RISK AND UNCERTAINTY,
2004 HURRICANE FORECASTS

1.1 Abstract

Hurricane landfall causes damage. This damage is neither unique across storms nor, in the short term, unpredictable. Property values along with hurricane characteristics (path and intensity) almost completely explain coastal property damage. Forecasts of these same characteristics explain more than half of eventual damages. Given these forecasts, this paper tests financial markets’ use of this information. Generalizing across all the storms of the 2004 season, I find financial markets to have a small but predictable response to hurricane forecasts. However this response fails to make efficient use of information in forming expectations of hurricane risk. Forecasts of hurricane path negatively impact U.S. insurers but only the certainty of hurricane landfall affects an index of the overall economy. Despite its role in explaining property losses, 2004 intensity forecasts significantly impacted neither insurer returns nor the overall economy.
1.2 Introduction

Multi-billion dollar damages are common when hurricanes come ashore. The magnitude of these losses coupled with their recent frequency and predictions of increasing severity (Emanuel 2005) make forecasting and mitigating hurricane losses a multi-trillion dollar industry.

The insurance industry in particular has played a large role in mitigation given exponentially increasing property values in hurricane prone states. Insurance, of any variety, spreads infrequent but large costs out over many periods. Replacing a large but infrequent cost with a series of frequent small payments, insurance provides customers with more predictable and constant finances. However, those selling insurance have assumed the underlying risk. Under traditional contracts, firms mitigate this risk by insuring a large number of customers whose risks (car theft, home fire, job loss) are not correlated.

Aggregation however is no help when risks are highly correlated across customers. In these cases, writing more insurance policies increases rather than decreases firms’ financial risks. Hurricanes are an archetypal example of highly correlated losses across numerous policies. A hurricane doesn’t strike a single home, but a multi-county area. Despite taking on substantial hurricane risk, I find the value of publicly traded insurers is only marginally impacted by hurricanes and hurricane
forecasts. While the larger U.S. economy is found to respond bluntly to hurricanes, ignoring forecasts of storms until landfall is certain, insurer stock returns are sensitive to forecasts over storm path.

Surprisingly empirical studies, like this one, over hurricane losses are rare. The scarcity is surprising not only due to the magnitude of past and predicted hurricane damages, but due to the wealth of publicly available information about storms and markets’ responses. Given concerns about existing markets’ ability to finance future catastrophic risk (Froot 1999), it would seem necessary to understand catastrophic events’ impact on insurance firms and the overall U.S. economy. A review of the handful of economic and financial studies of hurricanes points to a clear identification of available data over insurers’ exposure (Angbazo & Narayanan 1996, Born & Viscusi 2006) as well as information over storms and storm forecasts (Ewing, Hein & Kruse, 2006). However the conclusions of previous studies can’t be generalized due to their methodology. Dummy variables or event studies isolate individual storms, in essence ignoring the conclusions of meteorology (Southern 1979, Emanuel 1998) that hurricanes’ destructive potential can be generalized across 2 dimensions – path and intensity.

Unlike previous studies, I replace the dummy for a specific named storm (Andrew, Hugo, etc), with a continuous value to generalize the destructive potential across the entire 2004 hurricane season.
The benefits of this choice are five fold. First this continuous variable composed of storm path and intensity allows my conclusions to be generalized not just across all past storms or even past forecasts but to provide an estimate of future hurricane costs. Second by explaining 2004 insurance losses as a function of the National Hurricane Center (NHC)’s forecasts of hurricane path and intensity, I provide a financial confirmation of path and intensity’s ability to accurately predict hurricane related claims. These tests also provide a measure of the accuracy of NHC’s 72 hour forecasts in predicting hurricane destruction. Third by using the best real-time hurricane information available to the market, my conclusions are not hampered by the caveat of previous studies that cumulative effects of hurricanes may be different from those measured by event studies or dummy variables as markets may be pricing hurricane risk before storms ever come ashore. Instead this study serves as a test of market efficiency around hurricanes, finding that markets are indeed using forecasts in pricing insurer returns but also rejecting that all the valuable information put out by NHC is being used in this calculus. Fourth combining hurricane forecasts’ ability to predict losses and the ability of the same forecasts to predict insurer and market returns, I find hurricanes to have a larger impact on a market index then on an insurance index. However the effect on both insurer and market returns is found to be small both in terms of economic and statistical magnitude. Fifth generalizing across all the storms and

---

1 NHC provides its own exemplary quality control and tests of accuracy ([http://www.nhc.noaa.gov/verification/](http://www.nhc.noaa.gov/verification/)). Calibrating hurricane forecasts and realized characteristics, NHC’s mean 5 year (centered on 2004) error is 160 miles for path and 18 knots for intensity. However, unlike NHC this study is interested in calibrating forecasts accuracy in predicting hurricane financial losses.
forecasts of the 2004 season allows me to retest the typical hypotheses around stock response to hurricanes albeit with a much larger sample size.

While the three extensive databases used in this study are well known, this appears to be the first time they are used jointly. In the case of NHC’s hurricane forecasts, it is the first time for much of this information to be used at all in the economic or financial literature. Section 1.3 begins by introducing the data and information of interest.

Given the importance of model and variable specification to this study, Section 1.4 continues by testing the ability of 2004 hurricane characteristics and forecasts to explain state level insurance losses. Numerous tests of model specification are employed to show that combining information over path, intensity and exposure in a straightforward (but nonlinear) manner best explains insurance losses.

Section 1.5 then compares the differences in how hurricane forecasts explain insurance losses and the ability of the same information to explain indexes of insurer and market stock returns. By repeating the tests of section 1.4, section 1.5 shows that information (intensity forecasts) significant to explaining insurance losses is not used by the market in forming expectations of insurer or market value. While pricing of insurance returns is more refined than assumed by previous studies, I find the blunt instrument of hurricane landfall to do as well as forecasts in explaining the 50 basis point loss in the 2004 U.S. economy attributable to hurricanes. Section 1.5 concludes by graphically presenting the marginal effects of hurricane forecasts on three different
financial responses: insurance losses, insurance returns, and market returns. Assuming market pricing of hurricane risk to be a function of hurricane damage and finding this damage to be predictable in large degree, the three figures show potential improvement in the timing of market response to hurricanes.

In section 1.6, I differentiate the stock response of individual insurers using Zellner’s SUR. Imposing linear restrictions on this system of regressions allows tests of the uniformity of insurer responses to hurricanes conditional on risk. These tests find hurricanes to have significant impacts on insures with hurricane risk and impacts on insurers with no hurricane risk comparable to that of the larger economy. However the impacts on individual insures are found to be neither uniform in magnitude nor timing with information concerning hurricane risk being priced into different insurers returns at different points in storms’ synoptic life cycle. An explanation for both larger impacts on market returns is explored related to the idiosyncratic nature of insurance pricing.

After drawing conclusions in Section 1.7, Section 1.8 presents future research opportunities.

1.3. Data

In the Atlantic, hurricanes form during the warmer months of summer and early fall. While 2004 had a particularly long and active hurricane season with fourteen named storms, the only period to have a possible hurricane landfall in the United States was
The possibility of landfall is announced and updated by the National Hurricane Center (NHC) whenever there is a positive chance of a hurricane coming ashore in the US in the next 72 hours. These NHC forecasts come out roughly every four hours and detail not just the probability of hurricane landfall but the best path estimates of the storm over the next three days as well as forecasts of storm intensity.

NHC provides an archive of their forecasts to the public on line. These forecasts (as well as NHC’s own ex-post measures of prediction error) allow for an extremely fine grained analysis of hurricane path and intensity forecasts. Future analysis will take advantage of this precise data. This paper, however, uses a coarser estimate of hurricane path in modeling hurricane related losses and financial impacts. Reducing NHC’s forecasts to the probability of a hurricane coming ashore in a specific state and the corresponding hurricane windspeed provides both an improvement over previous studies and a reasonable approximation to hurricane path and intensity. These measures of path and intensity also allow NHC’s forecasts to be generalized across hurricanes and easily matched to publicly available data over insurer risk.

As part of the heavily regulated nature of insurance in the U.S., the National Association of Insurance Commissioners (NAIC) keeps records on aspects of insurers’ activities on a state by state basis. Instead of attempting to disaggregate this data, this study takes the opposite approach of aggregating hurricane forecasts to the state level. I combine NHC’s forecasts with information over the value of insured coastal property, and NAIC’s data on the value of specific insurers’ direct premiums. This combined
dataset of NHC forecasts and NAIC’s records will provide information to build the independent variables in this econometric study. To finalize the independent variables of Section IV, I also transform the absolute measure of insurer risk (billions of $ in insurance policies) into risk relative to firm size.

Information over firm size (market value) as well as the dependent variables of insurer and stock market returns comes from the Center for Research in Securities Prices (CRSP). In addition to the return of a value weighted index of all U.S. stocks, I collect the daily returns of 49 publicly traded firms whose primary business is insurance (CRSP code 6331). These 49 firms represent a cross section of insurers with respect to size and hurricane exposure. By matching these firms (and their subsidiaries) to the insurance policies written in hurricane prone states, I find 27 firms exposed to hurricane risk and accounting for almost a quarter of total premiums written.

This stock information shows 2004 was a good year for both the U.S. economy and the insurance industry, despite $42 billion in losses resulting from 8 hurricanes making landfall (some in multiple states). Insurer returns averaged 19% APR during 2004. However, the daily returns of insurers averaged -2% APR when facing some probability of hurricane landfall. A two sample t-test finds the difference in returns with and without short-term hurricane risk to be statistically as well as financially significant. Given the negative impact of hurricane information on returns, the remainder of the study focuses on identifying what aspects of hurricane information impact whom and to what degree.
1.4 2004 Hurricane Losses

To test the ability of path, intensity and forecasts to explain hurricane risk, this section models state-level insurance losses as a function of hurricane information. These tests find an extremely large economic value to meteorology and provide a measure of short-term hurricane forecast accuracy. Additionally, the empirically estimated functions suggest prescriptions for how hurricane info should be used by financial markets and functional forms used in Section IV and V to test markets’ observed response to hurricane forecasts.

2004’s hurricane landfalls can be broken down by state and storm. There were 14 combinations of state and named storms composing the year’s losses. These 14 combinations, along with storm landfalls that never occurred, were forecast by the NHC with hundreds of path and intensity forecasts between August 1st and October 15th. NHC puts out forecasts of storm path roughly every four hours, during periods when there is a chance of a hurricane making landfall in the U.S. within 72 hours. Averaging these forecasts into a daily forecast for each storm, state combination results in 208 predictions of possible hurricane path. While it is easy to see where 42 of these predictions could come from (14 events * three days), the source of the remaining forecasts are less evident. Because hurricane path is uncertain, NHC forecasts numerous paths with positive probability of landfall that do not occur within three days. This can happen when a storm changes direction either moving to hit another state or turning back out to sea. It can also happen when a storm takes longer than expected to
come ashore. In the first case, we have a storm/state path forecast associated with no hurricane landfall. In the second case there is a landfall associated with the path forecasts, albeit a few days later than originally expected by NHC.

Corresponding to these path forecasts are intensity forecasts. While the layman is familiar with the discrete 5 point scale of ‘hurricane category’, the meteorological literature (Southern 1979, Emanuel 2005) has identified wind speed cubed as a continuous measure of storm intensity. I begin the empirical analysis of hurricane forecasts on the financial sector by confirming the value of windspeed cubed in explaining hurricanes’ destructive force. Model 1 seeks to explain insurance claims as a function of storm intensity and coastal property values for the 14 storm state combinations that occurred in 2004. The results of this least squares regression are found in Table 1.1. Modeling hurricane damages as a function of landfall, coastal premiums, and wind speed cubed ($w^3$), explains 93% of insurance claims. Interestingly, allowing losses to be more complex functions (polynomial expansions) of wind speed does not increase the model’s explanatory power.

In addition to presenting the estimates of model 1, Table 1.1 presents the results of model 2, which explains both damages and eventual damages as a function of hurricane path and intensity as well as forecasts of these variables. Again to facilitate generalizability, dummy variables for states aren’t used but instead each state’s exposure is modeled continuously as the value of coastal insurance policies.
It may be important to emphasize again that the estimates of the accuracy of these 208 predictions include the storm state landfalls that occurred, the forecasts and corresponding damages that eventually occurred, as well as the forecasts that correspond to no damages. Using all ex-ante forecasts forces model 2 to estimate the accuracy of forecasts for both damages that did and didn’t occur. For state storm combinations that did occur $loss_k$ equals the value of insurance claims attributed to the specific hurricane. In the case of storms that were predicted to come ashore in a state, but never did, $loss_k = 0$. In both models 1 and 2, subscript $k$ represents each possible state storm combination.

Model 1: $loss_k = \beta_{\text{certain}}(ws^3*\text{coastalpremiums}_k*\text{landfall}_k)$

Model 2: $loss_k = ws^3*\text{coastalpremiums}_k*(\beta_{\text{predict}} \text{prob}_{\text{landfall}} + \beta_{\text{certain}} \text{landfall}_k)$

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{\text{certain}}$</td>
<td><strong>1.68E-8</strong></td>
<td><strong>1.68E-8</strong></td>
</tr>
<tr>
<td></td>
<td>(1.23E-9)</td>
<td>(1.49E-9)</td>
</tr>
<tr>
<td>$\beta_{\text{predict}}$</td>
<td>-</td>
<td><strong>3.25E-8</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.96E-9)</td>
</tr>
<tr>
<td># obs</td>
<td>14</td>
<td>208</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.93</td>
<td>.66</td>
</tr>
<tr>
<td>F</td>
<td>201</td>
<td>184</td>
</tr>
</tbody>
</table>

Table 1.1 Regressions of forecasts on eventual losses
T and F-tests find forecasts to be significant at >99% confidence level both individually and in conjunction with storm landfall in explaining hurricane losses.

The multiplicative nature of the variable composed of forecasts (and exposure) raises the possibility that a portion of this information contains all the explanatory power over hurricane losses. To test various forms of this hypothesis, J-tests are run on the predictive variable built according to theory against simplifications. The most extreme of these simplifications is the dummy variable approach used by other studies. Testing the simplest specification against progressive sophistication (exposure, exposure & path, exposure & path & intensity) J-tests reject simplified models ability to explain richer models again at >99% confidence levels. Conversely more sophisticated models can not be rejected against simpler models at any typical confidence levels. For illustrative purposes, I note that a variable that includes exposure and path forecasts has less than three quarters of the predictive power of model 2 which also includes intensity forecasts.

Table 1.2 uses the estimates of \( \text{loss}_k \) from model 2 to compare realized insurance claims against those predicted by hurricane information. While table 1.1’s parameter estimates and model statistics measure how insurance losses and hurricane information covary, table 1.2 compares the monetary value of realized and predicted losses for three randomly selected hurricanes. These storms represent a cross section of storm and state characteristics both on the day of hurricane landfall and three days prior. From the
table, it is evident that while not fully explaining the $42 billion in insurance losses from 2004 hurricanes, both hurricane characteristics and hurricane forecasts are extremely valuable in predicting insurance claims.

<table>
<thead>
<tr>
<th>Date</th>
<th>Hurricane</th>
<th>State</th>
<th>Prob of landfall</th>
<th>Wind speed</th>
<th>Realized insurance claims ($ millions)</th>
<th>Predicted insurance claims ($ millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 31st</td>
<td>Alex</td>
<td>nc</td>
<td>29</td>
<td>35</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Aug 3rd</td>
<td>Alex</td>
<td>nc</td>
<td>99</td>
<td>85</td>
<td>5</td>
<td>107</td>
</tr>
<tr>
<td>Sep 2nd</td>
<td>Frances</td>
<td>fl</td>
<td>26</td>
<td>125</td>
<td>4,110</td>
<td>3,203</td>
</tr>
<tr>
<td>Sep 5th</td>
<td>Frances</td>
<td>fl</td>
<td>99</td>
<td>95</td>
<td>4,110</td>
<td>2,765</td>
</tr>
<tr>
<td>Sep 12th</td>
<td>Ivan</td>
<td>fl</td>
<td>19</td>
<td>145</td>
<td>4,500</td>
<td>3,653</td>
</tr>
<tr>
<td>Sep 15th</td>
<td>Ivan</td>
<td>fl</td>
<td>99</td>
<td>120</td>
<td>4,500</td>
<td>5,572</td>
</tr>
</tbody>
</table>

Table 1.2 Realized and Predicted Hurricane Losses

In addition to their magnitude, a fundamental characteristic of hurricanes making them of continued interest to economists, meteorologists and coastal residents is the uncertainty surrounding hurricane predictions. However, as shown by these first results, uncertainty surrounding forecasts should not be confused for unpredictability of losses. Explicitly equating the 2 ideas has confounded previous studies (Lamb 1998) as well as citizen response to hurricane forecasts. Some financial studies have improved over this shortcoming by allowing markets to form expectations of hurricane landfall. The previous improvements however have been limited to specifying market
expectations with daily dummies beginning at an arbitrary point prior to a specific hurricane’s landfall.

Having identified a generalizable, publicly available predictor of hurricane damages, Section 2.5 measures both the importance of hurricane forecasts on market and insurance indexes and market efficiency in pricing hurricane risk. Specification tests will show markets to be more efficient than the simplified models of previous studies but not completely efficient in using NHC’s forecasts. The ability of this inefficiency to persist in the market will be explained by a combination of rare opportunities around hurricane forecasts for efficient traders and the relatively small magnitude of potential margins.

1.5. 2004 Stock Returns

The essence of the efficient market hypothesis is that stock prices immediately adjust to all relevant information over future losses or gains. In the current context of hurricanes the hypothesis maintains that the billions in losses predicted by NHC’s forecasts will immediately be priced by the market. Due to the peculiarities of insurance dynamics (Born & Viscusi 2006), it is not immediately clear whose returns should be affected by these valuable forecasts or if this information may even have a positive impact on some insurer returns.

The difficulty in prescribing how efficient financial markets should respond to hurricanes comes from the formation of two different expectations. The optimal
financial response becomes even murkier when attempting to prescribe insurer stock response to hurricanes.

The first expectation is over what new information is contained within a specific hurricane. Some general information is known about hurricane risk before a storm makes landfall or before NHC makes a short-term forecast of hurricane characteristics. At the extreme, we may safely assume that the market knew before the 2004 hurricane season began that hurricane risk in South Dakota was zero and that hurricane risk in Florida was greater than zero. The difficulty comes in prescribing how much information is known in general about hurricane risk and what new information comes from short-term hurricane forecasts.

Prior to tracking individual storms, meteorologists predicted a 70% chance of at least one hurricane making landfall in the U.S in 2004 (Gray & Klotzbach, 2004). The ambiguity and coarseness of this forecast points to the difficulty of long-term forecasts over hurricanes and weather in general. The increased information of short term forecasts of specific storms could potentially change expectations of insurer profitability in two different ways. If the market expected fewer or less intense hurricanes in 2004 then a negative insurer response to specific hurricanes may be expected. However, if the damages from a hurricane provide increased information (via Bayesian updating) about future hurricane losses we may expect markets to respond even more negatively to specific hurricane landfall.
The ambiguity over the extent of the negative insurer response to short-term forecasts is compounded by the second necessary expectation that markets must form. After updating their expectations over the likelihood of hurricane damages, markets must form an expectation of who holds hurricane risk. Within 2004, it is clear that insurers held a substantial hurricane risk as shown by $42 billion in insurance claims. However, the optimal inter-temporal expectation on insurer returns is less clear. In a competitive market, the $42 billion in 2004 claims may reasonably be expected to negatively impact insurers. However in the heavily regulated insurance industry, insurers may be able to recoup these losses through future premiums hikes. This possibility would essentially remove the negative impact of hurricane expectations from insurers and instead pass losses along to the overall economy through higher expected premiums.

These difficulties in forming optimal expectation operators have allowed efficient market explanations of empirical findings of negative (Ewing, Hein & Kruse 2006), positive (Lamb 1998) and insignificant (Angbazo & Narayanan 1996) insurer response to hurricanes. However in addition to explanations over hypothesized expectations, there is a methodological explanation for these studies’ mixed results. The mixed results of academic studies as to who is affected by hurricanes, how much, whether these impacts are positive or negative, and their timing are not surprising given the findings of section 1.3. Previous studies have omitted forecasts over hurricane losses, instead focusing on the market’s response to actual destruction caused by a
hurricane rather than its anticipated destruction. Allowing for the possibility that the market may use weather forecasts to incorporate expectations of hurricane damages into stock prices days before landfall, market returns on the day of hurricane landfall (or some arbitrary number of days before or after the event) would only capture the difference between expected and realized damages. This difference, in essence an error term, could just as easily be positive as negative based on forecasting accuracy and the markets’ use of that information. While extending the window to include longer and longer periods reduces this problem, it opens up new problems within the event study methodology (MacKinlay 1997) typically employed.

The continuous short-term expectation of 2004 hurricane risk defined in the previous section allows me to avoid both event studies and dummy variables all together. A simple OLS procedure will be used in this section to compare what hurricane information is being used in pricing an index of insurers and an index of the larger market. This information will also be compared against the short-term forecasts identified in table 1.1 as explaining two thirds of eventual hurricane losses.

Additionally the identified difficulties in prescribing whether insurers should be unaffected or positively or negatively impacted by hurricanes are by passed in this essay. Instead I first show empirically that both the market and insurer stock prices respond negatively to hurricane landfall. Then given the market’s empirical expectation that hurricane landfall is negative for both the overall economy and insurers, I am able
to test and reject that the market is efficiently using forecasts in timing their response to hurricanes.

Before regressing hurricane forecasts on stock prices, I modify the multiplicative variable, identified in models 1 and 2, to account for the continuously updated nature of stock returns. Because an efficient market prices stocks using all information on the day it is released, financial markets should not respond to forecast levels but changes in forecasts from day to day. Additionally, while the general economy is interested in the overall risk as defined in the previous models’ coastal premiums, insurers’ returns should respond to firm level risk defined as an estimate of a firm’s coastal insurance premiums\(^2\) divided by the firm’s market value. These two changes result in:

\[
\Delta \text{forecast}_i^t = \left( w_i^3 \times \text{prob}_i \times \text{premiums}_i^t / V_i^t \right) - \left( w_{i-1}^3 \times \text{prob}_{i-1} \times \text{premiums}_{i-1}^t / V_{i-1}^t \right) \quad (1)
\]

In equation (1) and throughout the remainder of the paper, days are indexed by the time subscript \(t\) and firms are indexed by superscript \(i\).

The nonlinearities introduced by changes in forecasts, rather than forecast levels, forces variable specification to be tested with nonnested model statistics (J or Cox

\[^2\] NAIC collects state-level data on insurance premiums but does not break this down between coastal and inland. I assume firms have a uniform distribution of policies across coastal and inland areas (a reasonable assumption given insurance regulations). To transform state level premiums to state coastal premiums I multiply total state premiums by coastal insurance/total insured property, per figure 6.
statistics) and/or information criteria (Akaike or Baysian). Replicating (1) for each of the k possible state storm combinations yields an econometric model of insurer returns:

\[ ret_i^t = \beta^* \sum_{k}^{208} \Delta\text{forecast}_{i,k} + \varepsilon_i \]

Estimating (2) using OLS is equivalent to assuming a homoskedastic equally weighted insurance index. While this assumption will be relaxed in the next section to model each firm’s individual stock response to hurricanes, the OLS estimate gives us the industry wide affect of forecasted hurricane losses.

Reporting the \( \hat{\beta} \) from equation (2) may be misleading. Unlike estimates of hurricane forecasts on insurance losses, the estimates of (2) are not robust to model specification. Adding variables changes both regression estimates and significance. Like in section 1.4, I will instead use tests over model specification to judge which information in \( \Delta\text{forecast} \) may be impacting insurer and market returns, which information may be extraneous, and which hurricane information may be entering into stock returns differently then theory suggests.

Table 1.3 presents alternative models explaining the daily return on indexes of insurers and the U.S. economy from Aug 1st to Oct 15th. Before drawing inferences from model differences, I will note the key similarities robust across all models.
Hurricanes are found to have a negative impact on both 2004 insurer and market returns. While exposure to billions in losses may be assumed to have a negative impact on firm value, this finding is far from universal in previous research. In addition to modeling concerns raised in this paper, other papers\textsuperscript{3} have explained mixed results by pointing out insurers ability to recoup current or past losses through future rate hikes. A well understood theoretical result, rarely pointed out in empirical studies, is the ability of risk mitigators to make higher profits in a world of higher risk, given the concave preferences of their customers. If increased current losses foretell increased future losses, insurer returns may reasonably be expected to respond positively to hurricane forecasts.

This line of reasoning explains the decision to model market returns as well as insurer returns. Billions in annual losses can’t make everyone better off. If insurers are expected to recoup their losses through higher premiums, then the market is absorbing these losses by paying higher premiums. I find the market to indeed respond negatively to hurricanes. This negative response could be the result of higher expected premiums or uninsured losses (i.e. economic losses as opposed to property damage). Interestingly the impact of hurricanes is larger both in economic and statistical terms for the market index than the insurance index.

However, for both the insurer and broader market indexes, hurricane forecasts do not go very far in explaining stock returns. Low explanatory power in other

hurricane studies may be taken as supporting the efficient market hypothesis. That is not the explanation here. While markets may be assumed to be more sophisticated than daily dummies in forming expectations of hurricane losses, they certainly have no better information than NHC’s forecasts. Instead the implication here is that hurricane information, while significant and negative both for insurer and market expectations, only explain between 1 and 10% of price volatility when the U.S. faces short-term probability of hurricane landfall.

\[
ret_t = \beta_0 + \beta_1 \sum_{k} D_k + \beta_2 \sum_{k} \Delta path_{t,k} + \beta_3 \sum_{k} \Delta forecast_{t,k} + \epsilon_t
\]

| Index of returns = f(E(losses)) = f(consultations) |

<table>
<thead>
<tr>
<th>Insurance Index</th>
<th>Market Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>\beta_0 Constant</td>
<td>\beta_1 Landfall Dummy</td>
</tr>
<tr>
<td>.26 (.31)</td>
<td>.31 (.31)</td>
</tr>
<tr>
<td>\beta_2 Path</td>
<td>\beta_3 Path &amp; Intensity</td>
</tr>
<tr>
<td>***-2.39 (.86)</td>
<td>***-8.15 (2.86)</td>
</tr>
<tr>
<td>***-1.98 (.87)</td>
<td>-1.95 (4.37)</td>
</tr>
<tr>
<td>***-2.40 (.87)</td>
<td></td>
</tr>
<tr>
<td>-.485 (2.88)</td>
<td></td>
</tr>
<tr>
<td>1.01 (1.04)</td>
<td>1.14 (1.04)</td>
</tr>
<tr>
<td>1.27 (1.04)</td>
<td>.77 (1.04)</td>
</tr>
<tr>
<td>1.14 (1)</td>
<td>.77 (1)</td>
</tr>
<tr>
<td>R²</td>
<td>.005 .012 .006 .10 .10 .08 .002</td>
</tr>
<tr>
<td>F-test (1 vs. 2)</td>
<td>J-test</td>
</tr>
<tr>
<td>***8.08</td>
<td>***2.81</td>
</tr>
<tr>
<td>**.92</td>
<td>-.06</td>
</tr>
<tr>
<td>1.22</td>
<td>.96</td>
</tr>
</tbody>
</table>

Table 1.3 Regress Hurricane Info on Stock Returns

Having explored the robust results of the models, I again focus on explanatory power (or model errors) to begin analysis of the differences in regression estimates.
The last two rows of Table 1.3 present model specification tests within three models of insurance returns and within four models of market returns. Clearly as these two sets of models use different independent variables, there are no cross-set tests of specification. However, comparisons reveal key differences in the information used by financial actors when pricing insurer and market hurricane risk. Both the information found to best explain market returns and insurer returns is significantly different from the theoretical model of equation (2).

Estimates of equation (2) are found in the final column of Table 1.3. Despite overwhelming evidence in favor of equation (2)’s ability to explain insurance losses\(^4\), this model does worse than a dummy variable \(\sum_k D_k\) in explaining market returns around hurricanes. J-tests fail to reject either a hurricane landfall dummy or a specification including hurricane path as appropriate for modeling market returns. A hybrid model containing both a hurricane dummy and forecasts over hurricane path is statistically indistinguishable from the simple dummy variable approach. Of the numerous theories proposed in empirical and theoretical studies of hurricanes’ impact on the economy, none can rationally explain this finding.

Given the empirical findings that market returns respond negatively to hurricane landfall (Table 1.3) and that hurricane forecasts do a good job of predicting damages from hurricane landfall (Table 1.1), efficient markets should use hurricane forecasts in pricing market returns. In contrast the apparent confounding of uncertainty and

\(^4\) See Section III
unpredictability found in hurricane studies appears to be a good model of financial behavior around hurricanes even though it isn’t a good model of hurricane losses.

Regressing hurricane forecasts on insurance returns finds this subsector of the economy to be responsive to NHC’s path estimates as well as hurricane landfall. The second model of the insurance index is a statistical improvement over both dummy variable approaches of previous studies and the theoretical model implied from Section 2.3, as shown by the results of the F and J tests reported in table 1.3. The implication that financial markets can price some portion of insurer risk as a function of the probability of hurricane losses before a hurricane comes ashore is reassuring. Even finding landfall to include more information than forecasts alone is not unsettling given errors in NHC forecasts. The surprise of insurer stock response to hurricanes is that intensity estimates are not used. This is true despite these estimates simultaneously composing roughly a quarter of the ability of forecasts to explain eventual hurricane related insurance claims.

Figures 1.1-1.3, summarize the findings of this and the previous section. Presenting the marginal effects of path and intensity forecasts on 1) U.S. returns, 2) Insurer returns, and 3) Hurricane damages, the figures graphically contrast the impact of 2004’s hurricane forecasts on the economy. Pricing and trading the market (insurance) index according to figure 1.3, an investor could make economic profits equal to the difference in figures 1.3 and figure 1.1 (1.2).
1.6 Individual Insurer Returns

Differences in financial responses to hurricanes and hurricane forecasts are not limited to loss, industry or economy wide indexes. All the studies cited have raised the possibility of differences in insurer stock responses to a hurricane based on firm characteristics.

Dividing the 49 firms along any characteristic (size, exposure, risk mitigation measures) we could create indexes of firms and repeat the techniques of the previous section. These estimates would be unbiased, but wouldn’t take advantage of efficiency gains possible given multiple observations (one for each index) on any given day. Efficiency improvements grow in proportion to the similarity of non-modeled factors’ impact on each group. Given 49 firms operating within the same country, industry and regulatory environment it is very likely that announcements and market shocks will cause similar responses.

This section uses Zellner’s SUR to exploit potential efficiencies while simultaneously facilitating hypothesis testing across differences in individual firms’ responses to hurricanes. The system of equations, presented in (3), is based on the best model of section IV and estimates individual insurer response to hurricane forecasts and landfalls.
\[ ret_i = \beta_i + \beta_{id} \sum_{k}^{14} D_k + \beta_{ip} \sum_{k}^{208} \Delta path_{t,k} + \beta_{im} ret_m + \epsilon_i \]  

(3)

\[
\begin{align*}
\mathbf{\epsilon}_i = \begin{pmatrix}
\epsilon_{i1} \\
\vdots \\
\epsilon_{iT}
\end{pmatrix} \\
\mathbf{ret}_i = \begin{pmatrix}
ret_{i1} \\
\vdots \\
ret_{iT}
\end{pmatrix} \\
\mathbf{ret}_m = \begin{pmatrix}
ret_{m1} \\
\vdots \\
ret_{mT}
\end{pmatrix}
\end{align*}
\]

where:

\[ E(\mathbf{\epsilon}_i) = 0 \]

\[ \text{cov}(\mathbf{\epsilon}_i, \mathbf{\epsilon}_j) = \begin{cases} 
\sigma_{ii}^2 = \sigma_{ii}^2 & i = j, t = s \\
\sigma_{ij}^2 & i \neq j, t = s \\
0 & t \neq s 
\end{cases} \]

In addition to estimating a separate hurricane response for each firm, equation (3) includes an estimate, \( \hat{\beta}_{im} \), of each firm’s sensitivity to changes in the overall markets’ rate of return.

Comparing OLS estimates of (3) to GLS estimates that account for intra-temporal correlation of errors finds significant efficiency gains to GLS. A Durbin-Watson test finds no autocorrelation, \( \sigma_{ts} = 0 \), supporting the final assumption of the model. Accepting the model’s specification, the remainder of this section focuses on estimates and hypothesis testing.

---

5 Breusch-Pagan test rejects \( T \sum_{i=2}^{n} \sum_{j=1}^{i-1} r_{ij}^2 = 0 \) at > 95% confidence. Where \( r_{ij}^2 \) is the residual correlation coefficient of firm \( i \) and firm \( j \).
With 49 insurees and 4 variables, GLS yields almost 200 parameter estimates. Figures 1.4 and 1.5 show the distribution of insurer stock responses to hurricane landfall and path. The direct inference is that the average impact of hurricane forecasts estimated for an insurance index contains a fair amount of variability across individual insurers. Significant estimates of insurer stock response to path forecasts ranges from extremely negative to slightly positive. We can attempt to explain this range of estimates through a cross-sectional analysis. Because meaningful firm characteristics (reinsurance, loss reserves, ceding premiums, Best’s credit rating, market capitalization) do not vary over the estimation period, we can regress these characteristics on each firm’s coefficient estimates. These characteristics are not found to be significant in explaining the variance in parameter estimates. While these characteristics include risk mitigation strategies that limit a firm’s loss exposure on the books, in both this analysis and at least one previous study (Angbazo & Narayanan 1996) little impact on the market’s perception of insurers’ hurricane risk is shown.

This finding of risk-mitigation playing no role in explaining stock response to hurricanes can be taken and tested at its logical extreme due to geographic differences in insurers. Of the 49 insurers, 22 wrote no property policies in hurricane prone states. A common hypothesis in the financial literature maintains that these firms may still be negatively impacted due to ‘contagion’ affecting solvency across the industry regardless of exposure. A weak test of contagion holds that the overall impact of hurricane landfall on unexposed insurers is no different from the impact on the broader economy.
This can be tested through the linear restriction \( \sum_{j=1}^{22} \hat{\beta}_{jd} = 0 \) on the SUR estimation of (3). This restriction can not be rejected at any reasonable level (\( F=.08 \)). A stronger version of the test is whether each non-exposed insurer was significantly affected by hurricane landfall, \( \hat{\beta}_{1d} = \hat{\beta}_{2d} = \ldots = \hat{\beta}_{jd} \ldots = \hat{\beta}_{22d} = 0 \). An F-test (1.41) fails to reject the stronger restriction. So while risk-mitigation can not explain cross sectional differences in insurer stock prices, the market did not use hurricane information in pricing unexposed insurers, apparently taking into account that almost half of insurers did not write property policies in states with 2004 hurricane risk.

Appealing to the same linear restrictions we can test two other hypotheses. Tests of the same restrictions on \( \hat{\beta}_d \) for the 27 firms with hurricane risk, reject the restrictions for both the weak (\( F=3.43 \)) and strong (\( F=2.66 \)) hypothesis of no impact of hurricane landfall. The implication being that hurricane landfall has both a significant and consistently negative impact across exposed insurers. Imposing linear restrictions on the 27 exposed insurers, we can also test the significance and consistency of the effect of hurricane forecasts. In this case, I find forecasts significantly (\( F=9.50 \)) impact individual insurers’ stock returns. However, the sign of the impact is not consistent across insurers (\( F=.04 \)).

Together with the findings of Section IV, these hypothesis tests find hurricanes to have a significant impact on insurers. While the impact of hurricane landfall is

---

6 Where \( j \) represents the 22 insurance firms with no hurricane exposure.
negative and near universal, markets are also found to use hurricane forecasts in pricing insurer returns. Overall markets expect hurricane forecasts to negatively impact insurers, however a few insurers are found to respond positively to path forecasts. Despite significant differences in a few insurers’ stock response to hurricane path, firm characteristics are not found to be correlated to these differences.

1.7 Conclusions

The last few years have seen record losses due to hurricanes and rising concern over the ability of insurance to cope with these losses. Despite this concern there has been surprisingly little empirical research over hurricanes’ impact on financial markets.

While data is certainly not lacking in this field, finding an appropriate methodology to model expectations of hurricane damages has been a problem in financial research. This paper uses real-time publicly available forecasts to create a single continuous measure of short-term hurricane risk. This measure of path and intensity is generalized across all of 2004’s storms and regressed on property damage claims resulting from hurricane damages. While the measure is found to explain hurricane losses to a high degree, it is statistically rejected in favor of simpler specifications for explaining stock price response to hurricane forecasts.

Explaining returns as a function of exposure as well as hurricane path and intensity allows this paper to test the market’s use of forecasts in setting insurer and market indexes. I find information highly valuable in explaining hurricane losses not to
be used in pricing market risk around hurricanes. Both indexes of insurers and the overall economy respond negatively to damages from hurricane landfall. Given that these changes in stock price are significant and predictable, any information that forecasts these damages should immediately be used to price hurricane risk before hurricane landfall. However I find that markets are not using hurricane intensity forecasts in pricing the insurance index and are using neither intensity nor path forecasts in pricing the overall market index. Additionally, I find significant differences across insurers, independent of risk mitigation strategies.

2004 hurricane path and intensity forecasts allow this paper (and future research) to not only test the information set used by markets in pricing risk, but allows testing of existing financial hypothesis, albeit with a much larger sample. Generalizing across all the hurricanes and short-term hurricane forecasts of 2004, I find the overall U.S. economy to be impacted by landfall, unexposed insurers to be affected similarly, and insurers writing policies in hurricane states to be impacted both by landfall and forecasts of hurricane path.

Despite statistical significance over numerous metrics and corresponding to $42 billion in insured losses, hurricane information is found to have only marginal explanatory power over stock returns. A slightly larger impact on the overall economy in contrast to insurers, points to an ability of existing markets to finance hurricane risk and pass increased costs along to the insured or larger economy. Whether this is accomplished through reinsurance or insurers’ ability to recoup current losses through
future increased premiums is not clearly identified. However, the insignificant cross-sectional estimates of insurer characteristics (including reinsurance) on firm response to hurricanes points to inter-temporal premium pricing as the likely answer.

1.8 Caveats & Future Research

Two potential expansions of the data used in this study would provide potential improvements and allow for more refined hypothesis tests around market response to hurricanes and hurricane forecasts.

First, this study has shown that hurricane damages are generalizable across storms using forecasts over path and intensity. I have also shown that market responses to hurricanes can be modeled on more than a storm by storm basis. While this study has used forecast and market data from an active hurricane season, there is no reason that the same methodology could not be expanded across numerous years. Expanding the dataset to include multiple years of hurricane forecasts and insurer exposure would allow for more accurate estimation of model parameters as well as allowing for tests of models’ out of sample prediction accuracy. The value of testing out of sample predictive power is evident. What is less evident is increased data’s potential for allowing less restrictive modeling assumptions over how markets use information around hurricanes. Parametric assumptions are necessary in the current study due to small sample size. Luckily the meteorological and financial literature provides insight into a correct parameterization. However having found that this correct specification is
not being used by financial markets, there are an infinite number of incorrect uses of hurricane information that markets could use to approximate the correct specification. Adding more years to the study could allow for alternative methods (nonparametrics, polynomial expansions) of estimating market use of hurricane forecasts in pricing hurricane risk. In essence these alternative methods may serve as a substitute for the non-nested hypothesis tests of sections 1.4 and 1.5. Potentially these methods would allow for one model to be rejected over another as was done with the J-tests for the insurance index but was not possible with a J-test for the market or a Cox test for the SUR model of individual insurers.

The second potential improvement from increased data stems from the geographic accuracy of the measure of value at risk. As explained in section 1.3, NAIC only has state level insurance data. Transforming state level insurance into value at risk in this study requires 1 step and 2 assumptions. The step is to multiply each firm’s state level insurance premiums by the percentage of that state’s total premiums written on the coast, per figure 6. The 2 assumptions associated with this transformation are 1) property values are uniformly distributed across a state’s coast 2) each insurer’s policies are evenly distributed between coastal and inland areas. Despite the obvious inaccuracy of the first assumption, it is shown in table 1 to allow an accurate estimate of hurricane losses. The second assumption can not be tested with the current data and instead rests on state insurance commissioners requiring companies who write inland property insurance to also offer coastal property insurance. To test the second assumption and
potentially improve estimates of insurer stock returns, county level data could be used in place of state level data. While I currently do not have access to this data due to its proprietary nature, it clearly exists (Watson, Johnson, Simons, 2004) and may allow even larger improvements in measuring and pricing hurricane risk.
References


Figure 1.1 Market Index Response to Δ forecast.
Figure 1.2 Insurance Index Response to $\Delta$ forecast.
Figure 1.3 Insurance Losses Response to $\Delta$ forecast.
Figure 1.4 Histogram of $\hat{\beta}$, impact of path forecasts on insurer returns.
Figure 1.5 Histogram of $\hat{\beta}_{id}$, impact of landfall on insurer returns.
Figure 1.6  VALUE OF 2004 INSURED COASTAL PROPERTIES VULNERABLE TO HURRICANES BY STATE

<table>
<thead>
<tr>
<th>State</th>
<th>Coastal (billions)</th>
<th>Total exposure (2) (billions)</th>
<th>Coastal as a percent of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida</td>
<td>$1,937.4</td>
<td>$2,443.5</td>
<td>79%</td>
</tr>
<tr>
<td>New York</td>
<td>1,901.6</td>
<td>3,123.6</td>
<td>61%</td>
</tr>
<tr>
<td>Texas</td>
<td>740.0</td>
<td>2,895.3</td>
<td>26%</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>662.4</td>
<td>1,223.0</td>
<td>54%</td>
</tr>
<tr>
<td>New Jersey</td>
<td>505.8</td>
<td>1,504.8</td>
<td>34%</td>
</tr>
<tr>
<td>Connecticut</td>
<td>404.9</td>
<td>641.3</td>
<td>63%</td>
</tr>
<tr>
<td>Louisiana</td>
<td>209.3</td>
<td>551.7</td>
<td>38%</td>
</tr>
<tr>
<td>South Carolina</td>
<td>148.8</td>
<td>581.2</td>
<td>26%</td>
</tr>
<tr>
<td>Virginia</td>
<td>129.7</td>
<td>1,140.2</td>
<td>11%</td>
</tr>
<tr>
<td>Maine</td>
<td>117.2</td>
<td>202.4</td>
<td>58%</td>
</tr>
<tr>
<td>North Carolina</td>
<td>105.3</td>
<td>1,189.3</td>
<td>9%</td>
</tr>
<tr>
<td>Alabama</td>
<td>75.9</td>
<td>631.3</td>
<td>12%</td>
</tr>
<tr>
<td>Georgia</td>
<td>73.0</td>
<td>1,235.7</td>
<td>6%</td>
</tr>
<tr>
<td>Delaware</td>
<td>46.4</td>
<td>140.1</td>
<td>33%</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>45.6</td>
<td>196.0</td>
<td>23%</td>
</tr>
<tr>
<td>Mississippi</td>
<td>44.7</td>
<td>331.4</td>
<td>13%</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>43.8</td>
<td>156.6</td>
<td>28%</td>
</tr>
<tr>
<td>Maryland</td>
<td>12.1</td>
<td>853.6</td>
<td>1%</td>
</tr>
<tr>
<td>Coastal states</td>
<td>6,863.0</td>
<td>19,041.1</td>
<td>36%</td>
</tr>
</tbody>
</table>

(1) Includes residential and commercial properties. Ranked by value of insured coastal property.
(2) Exposure is the total amount of insured property in the state.
Source: AIR Worldwide.
CHAPTER 2

KNOW WHEN TO HOLD’EM

2.1. Abstract

Professional poker players are not dynamic optimizers. The surprise is that their behavior is overly conservative. Suboptimal decisions may result from either loss aversion or heuristics that poorly approximate dynamics in the credit constrained environment of high-stakes tournament poker. Additionally, I find that despite behavior consistent with players’ comments that they can ‘read’ their opponents, players’ expectations of having the best hand are no more accurate than simple statistical models that assume opponents’ cards are randomly distributed. The implication of professionals in a high-stakes competition confusing noise for information and making suboptimal choices is explored.
2.2 Introduction

High-Stakes Poker is as an ideal environment to study the optimizing choices of professionals. Avoiding confounding explanations for behavior found in many other real world problems, poker preserves essential elements of decision making under risk and uncertainty.

In a televised variant of poker, professional gamblers invest thousands of dollars to enter the zero-sum game of tournament poker where they face repeated decisions to fold or bet in an attempt to win millions of dollars. At each decision, players know the strength of their own cards but hold incomplete information as to their opponents’ cards and thus their true probabilities of winning. Similar uncertainty characterizes much of the risky decision making of interest in the real world. Whether we are interested in the entry/exit decisions of a firm in a market or the decision to buy/sell a stock, firms rarely face complete information about the risks they face. Instead they must separate noise from valuable information in forming expectations of the future and using these expectations to maximize profit.

Empirical studies of risky decisions made in these circumstances have documented systematic differences from expected profit maximization. Laboratory subjects are both overconfident when predicting uncertain events (Kahneman & Lovallo 1993) and make
systematically more conservative choices than neoclassical economics predicts. This conservative behavior has been found both when facing certain risk - (Kahneman, Knetsch & Thaler 1991) and uncertainty (Ellsberg 1961, Hogarth & Kunreuther 1989). The divergence from profit maximization is intensified when - like in tournament poker or other dynamic optimizations - the risky decision doesn’t lead to an immediate cash payout but to a lottery of outcomes (Slten, Sadrih & Abbink 1999).

In response to these experimental findings, many note Friedmen’s (1953) hypothesis that the greater opportunities for profit and learning found in competitive economic environments would correct or drive behavior to the neoclassical model. While the argument is intuitively appealing\(^7\), it remains to a large degree just a maintained hypothesis due to the difficulties of: 1) prescribing the optimal choice for a real firm 2) collecting data, and 3) empirically testing the outcomes of firms’ decisions. Recruiting professionals as subjects to simplify the problem, a few laboratory studies (Kagel & Levin 1986, List 2004) have found experience to help overcome some behavior, but when facing risk professionals’ decisions remain similar to - and sometimes worse than – amateurs’ (Hogarth & Kunreuther 1989, Haigh & List 2005).

Alternatively, rational expectations and profit maximization can be tested in simplified, highly-visible economic competitions. Following this approach, this paper opens another arena, in addition to stock analysis (Locke & Mann 2000, Engelberg, Sasseville, & Williams 2007) and professional sports (Romer 2006) where we find

\(^7\) De Long, Shleifer, Summers, Waldmann (1991) provide a counter argument for noise traders not just surviving but thriving in financial markets.
examples of professionals’ decisions diverging significantly from optimization. While much stock analysis suffers from issues of agency and complexity that confounds theoretical prescriptions, many specific sports decisions allow for theoretical models.

Both the simplification allowing for decision modeling in sports as well as the risk, uncertainty and dynamics of interest to economics are found in spades in poker. Agency issues are avoided because the firm is comprised of a single player, facilitating modeling the objective function as maximization of expected profit. Additionally the information (and noise) going into players’ expectations is limited to the cards and opponents around the table. For these reasons, economics has a history of interest in poker. Von Neuman and Morganstern presented expected utility theory using examples drawn from poker. Similarly, Nash chose to present his advances in game theory using poker.\(^8\)

This is not to say poker is a simple game or that optimization is trivial in this setting. Complexity - in the form of incomplete information, dynamic considerations, and, in the case of high stakes tournaments, credit constraints – provides abundant challenges to the casual and professional gambler alike.

Focusing on a single class of decisions within high-stakes tournament poker, I test how well professional gamblers handle these challenges. Professional gamblers exhibit the same tendencies of overconfidence and overly conservative decisions as subjects in the laboratory. This suboptimal behavior comes from both players’ mistaken belief that

---

8 Both the lexicon of statistics (Monte Carlo) and finance (blue chip stocks, call) also reflect those fields interest in the game from as far back as the Bernoullis in 1700
they can ‘read’ their opponents (confusing noise for information in the behavior of their opponent) and their need for a risk premium even after paying up to $25,000 to enter a zero sum game. In the three years of decisions analyzed, players’ won half as much as they would had they followed either a dynamic or even a simple static EV (expected value) maximization rule.

These results reject hypotheses maintained by both Friedman and the similarly seminal poker player/theorist Doyle Brunson. According to Brunson, the complexity of separating information from noise in a dynamic setting “…is why a computer will never play top-notch poker”. While a computer has yet to play top notch poker, both the ability of professionals to separate information from noise and maximize their odds of success by ignoring human foibles of risk, loss and/or ambiguity aversion are called into question.

Section 2.3 begins the exposition by summarizing the rules of high-stakes poker relevant to the paper. I continue by limiting the focus to the final decision within a ‘hand’ of poker to create a prescriptive model of when a player should bet or fold. As will be shown, this repeated decision proves to be adequately rich to prove computationally difficult. However the investment pays off in a theoretical model of the optimal choice of a risk-neutral poker player. Focusing on just the final decision facilitates this modeling by avoiding both the option value present in earlier decisions and strategic behavior (i.e. bluffing).

---

9 Interestingly in the first Man-Machine Poker Championship an EV maximizing computer program did well against 2 top ranked players, albeit in a slightly simplified game (Limit heads up) from the one analyzed here. [www.cs.ualberta.ca](http://www.cs.ualberta.ca)
Section 2.4 describes a rich new data set, provided by the recent explosion of interest in poker. While poker has interested economists for at least 60 years and statisticians for hundreds of years, recent developments make this one of the first data sets available for empirical tests of professional behavior. The popularity, risks and rewards of high-stakes poker have created a modern cadre of professional gamblers not unlike professional options traders. These gamblers have decades of experience in a zero-sum game where they put up thousands of dollars for the potential to win hundreds of thousands to millions in rewards. The decisions of these professional have only recently become publicly available through a number of broadcasts of high-stakes tournaments, including the World Poker Tour used in this study. These broadcasts provide the relevant economic variables of interest of both costs of betting and potential benefits. While this information has always been available to observers around a poker table, the broadcasts additionally provide a view of the cards a player holds. This new combination of information allows section 2.4 to empirically estimate the expected value of betting. Section 2.4 concludes by econometrically estimating the parameters of a representative player’s objective function and testing for attendance to some of the behavioral anomalies identified in the laboratory.

In addition to broadcasting a player’s own cards, televised poker provides the viewer with information about the opponents’ cards as well. Taking advantage of this information, Section 2.5 tests the observed outcomes of players’ decisions against what would have happened had the player followed the prescriptions of the theoretical or
empirical models. The results confirm players to be overly conservative. Players are shown to win and lose significantly less than either static or dynamic theoretical models. The tests of Section 2.5 additionally show that the signals players are using in forming their expectations of success contain as much noise as information. Players’ actual decisions result in no difference in outcomes from conservative empirical models that only use a player’s own cards in estimating her probability of winning.

Section 2.6 concludes by discussing the role of players’ divergence from theory both for their odds of winning the tournament and the results’ broader implications.

2.3 Optimization in Poker

A. Description of Tournament

Of the many varieties of poker, No Limit Texas Hold’Em (Hold’Em) has captured the fascination of the American public. Tournaments of this poker variant are frequently broadcast on a host of television channels. The same characteristics that make these tournaments interesting to high-stakes gamblers and the public, namely professional participants competing in a high-risk environment for millions of dollars, also makes the game ideal for testing optimizing behavior.
A single ‘hand’ of Hold’em consists of four stages, each followed by a round of betting.\textsuperscript{10} The bets a player makes in the first three stages produce no immediate reward. Instead these bets are investments which both allow a player to continue to the last stage and influence the distribution of both rewards and odds of winning. Players who remain in the hand until the fourth (river) round of betting must then make a terminal decision to call (meet the fourth stage investment cost) or fold (exit the hand forfeiting all previous bets and opportunity to the reward at the end of the hand). After this fourth and final round of betting, those players who remain in the hand, proceed to the showdown where they compare their cards. The player with the highest five card combination\textsuperscript{11} wins the pot (sum of all players’ previous bets).

Virtually all data on professional gamblers’ decisions come from tournaments. Tournaments consist of multiple hands of Hold’em, where all players pay an entry fee before the first hand. In the first hand, all players begin with an identical number of chips with which they can bet. Players begin subsequent hands with their initial chips minus (plus) their bets lost (pots won) from previous hands. This redistribution of chips continues over multiple hands with players leaving when they run out of chips. In televised tournaments, a player is neither allowed to buy-back into the tournament nor exchange her chips for cash. The tournament ends when only one player is left with all the chips at the table, at which point the prizes are distributed.

\textsuperscript{10} The four stages are usually referred to by colloquisms of disputed origin: 1) hole 2) flop 3) turn 4) river - See Appendix A for game tree
\textsuperscript{11} Hold’em’s five card combination ranks are standard and shared by virtually all forms of poker
The repeated hands and credit constraints of tournaments add an additional dimension to the dynamic optimization. Players should no longer maximize the Expected Value (EV) of a hand. Instead they are attempting to maximize the probability of winning the tournament. One of Brunson’s contemporaries more inclined to the language of statistics and mathematics explains the difference in terms of EV:

"In this book, EV is mentioned frequently because it turns out that in a tournament it is not always right to choose the play with the slightly higher EV. This is because the higher EV bet may be more likely to lose...Assuming that you have enough money to withstand short-term fluctuations, it is always better to choose the bet with the higher EV. But if you do not have that cushion, it may well be right to choose the slightly smaller EV if that bet will win more often, especially if going broke keeps you from making more positive EV bets."

_Sklansky Tournament Poker for Advanced Players pg 17_

Sklansky’s advice implies that advanced players both understand profit maximization and are so inclined to static maximization that they must be dissuaded when in tournaments from making bets that maximize their chips at the end of the hand. This implication is supported in reading Sklansky’s book or Brunson’s (1978) where stochastic maximization and expected value are presented precisely and better than in many financial analyses or economic texts. Sklansky focuses the remainder of his book on presenting heuristics to aid players in going from static or single hand optimization to dynamic or tournament optimization.

Given this attention to EV in poker, I develop models of both dynamic and static maximization for a single class of poker decisions. These theoretical models are then
tested empirically to see how closely the observed decisions and outcomes of professional gamblers approximate optimization.

B. Value of Fold/Call in Final (River) Round of Betting

This section proposes two theoretical models of poker where players make a single decision within each hand to maximize either their expected stack at the end of the hand or probability of winning the tournament. The model precisely represents a class of choices – namely the final (River) decision of a player to fold or call.

In the River round of betting, players observe a number of economic variables pertinent to their decision to fold/call. Remaining players have matched each other’s bets in the previous three rounds. The sum of these bets make up the pot. The players also observe the amount their opponent has bet in this round\(^\text{12}\). In addition to increasing the pot, bet is the amount a player must call to proceed to the showdown. A player uses her current endowment of chips or stack to call the bet. However if her stack is smaller than the bet, she is always allowed to call by going all in (investing her entire stack). In this case instead of winning the pot plus 2 times the bet, having the best cards only entitles her to the pot plus 2 times her stack. The decision tree is presented in figure 2.1:

\(^{12}\) I ignore the case where bet=0 making players’ choice to stay in the game trivial.
From the tree, we can see that in addition to the observables which determine the payouts, the player must form an expectation of her probability of winning - $E(\text{prob})$. Forming this expectation is a key skill in the game. Its importance is pointed out by virtually every player and commentator and addressed at length in poker advice books. A basic expectation only uses the information known from a player’s two hole cards and the five community cards.\textsuperscript{13} This probability can be calculated from basic statistics assuming an opponent’s hole cards are random. I’ll refer to this as the naïve expectation. While in their commentaries and books professional gamblers begin with this naïve expectation, they propose updating it with a number of complex and situation specific rules. These rules boil down to differentiating information from noise in the behavior of their opponent. These conditional expectations have a striking resemblance to professional traders’ expectations of stock price, which begin with the Capitol Asset Pricing Model (CAPM) and are updated with myriad situation specific rules. While

\begin{figure}
\centering
\includegraphics[width=\textwidth]{decision_tree.png}
\caption{Decision Tree – Poker final decision}
\end{figure}

\textsuperscript{13} See Appendix A
both players and traders may believe these conditional probabilities contain valuable information, it comes at the cost of uncertainty. Additionally, just as traders’ expectations have been found to contain as much noise as information, players’ conditional expectations fail to improve their success as will be shown in Sections 2.3 and 2.4. The formulation, estimation and testing of players’ E(Prob) will be left for later.

Once the expectation is formed, the optimal static choice, $x=[0,1]$, that maximizes chips is simple: call ($x=1$) if the expected value of going to the showdown is greater than the value of folding. Equation (1) transforms the decision tree of figure 2.1 into the decision rule to call.

$$ [prob \cdot (\min\{2 \cdot stack, bet + stack\} + pot) + (1 - prob) \cdot \max\{0, stack - bet\}] \geq stack \quad (1) $$

This rule corresponds both to optimization in cash games\(^{14}\) and Sklansky’s EV maximization and is therefore a candidate for players’ objective function. However, dynamically, the value of folding/calling in tournaments is slightly more complex. According to the Bellman principle, $V_t(x_i|s)$ is equal to the next period’s expected value, $E[V_{t+1}(s)]$. The value of folding/betting in this period is conditioned upon the realization of the state variables ($prob, pot, bet, stack$) while next period’s value is conditional on the player’s new $stack$ and the following periods’ realizations of the

\(^{14}\) Unlike tournaments, cash games allow players to leave or enter at the beginning of each hand. The flow of players in and out of a game makes cash games popular in casinos but non-existent on television – limiting empirical data on players’ decisions.
other three stochastic variables. This value function also doubles as the expectation of winning the tournament, when I set the final reward equal to one if player $i$ wins the tournament, $\text{Stack}^i_{T+1} = \sum_{j=i} \text{Stack}^j_{T+1}$, and zero otherwise, where $j$ indexes each of the players around the table.

Reducing subsequent hands’ choices to a single decision to fold or call, allows figure 2.1 to be solved as a Markov process with only two minor changes. First, a parameter must be subtracted from each of the payouts representing the average sunk investments required by the game (antes, blinds, and previous bets) before a player is allowed to make her final decision in the next hand. Second, while $\text{Stack}$ continues to evolve according to the process of figure 2.1, we no longer consider $\text{pot}$, $\text{bet}$, and $\text{prob}$ to be deterministic. Instead these observed values are taken as realizations of stochastic states.

As presented in section 2.3 describing the empirical data, it turns out that the three stochastic state variables and the average sunk investment parameter are functions of the number of players at the table. As the tournament progresses, players are eliminated and chips becomes more concentrated. Costs and benefits increase not just in nominal terms but even relative to remaining players’ stacks. For this reason, I present equation (2), the Bellman equation for a risk neutral dynamic optimizer, with a fifth and final state variable $n$ (number of players). In the context of states conditioned on the number of players remaining in the tournament, realizations of the three discrete
stochastic state variables can be considered as the step direction of a shock (e.g. smaller, small, average, large, larger) with the number of players determining step length.

\[
V_t(s) = \max_{x \in [0,1]} \left\{ \sum_{s' \in S} P(s'|s,x)V_{t+1}(s') \right\} \text{ where } s \in S = \left( \text{prob } \otimes \text{pot } \otimes \text{bet } \otimes \text{stack } \otimes n \right)
\]  

(2)

In this dynamic maximization, a player chooses to fold or call, \( x \in [0,1] \), to affect the transition probability matrix, \( \sum_{s \in S} P(s'|s,x) \), such that her probability of winning the tournament, \( V_t() \), given her current realization of the five state variables, \( s \), is maximized. Because a player receives no reward, except once the entire tournament is over, the current value of her hand is completely determined by how her current decision affects the distribution of state variables, \( s' \), she may encounter in the next period, \( t+1 \). The universe of possible states is defined by the Kroenecker product of the first four states and conditioned on the fifth state, \( n \), the number of players who remain around the table.

A professional gambler may be hard pressed to identify her own behavior in this formulation of the bellman equation. But it is clear to both her and us that her stack in the next hand and subsequently her odds of winning the tournament are determined by her choice to call or fold as a function of her probability of winning the pot and the amount of her stack she must bet to stay in the hand.
The only remaining difficulty in solving the optimization lies in filling in the transition probability matrix. Because the first three state variables are independent across hands, the probability of any realization of these variables is conditional only on the non-parameterized empirical distribution described in section 2.3. What is highly conditional on the current state and decision to fold or call is a player’s stack in the next hand. This portion of the transition probability matrix is defined by figure 2.1. The transition of the final state variable, n, assumes a ten-percent chance of losing a player each round.

Given the option values and strategic behavior (i.e. bluffing) present at bets previous to the final decision, it is reiterated equation 2 should not be taken as a model of every decision in tournament poker. Instead (2) is a good representation of a specific decision within tournament poker – the choice to fold or call in the final (river) round of betting. By focusing on this single narrow decision within the game, the models of this section avoid the strategic behavior found in previous stages of the ‘hand’ and eliminate multiperiod dependence of states on previous states and actions. While avoiding these confounding explanations, Section 2.3 uses the static EV model to estimate the representative player’s revealed objective function. Section 2.4 uses the prescriptions of both static and dynamic models together with the empirical data to compare the outcomes of observed decisions with theoretical benchmarks.

15 The entire matlab code for the optimization is available upon request
16 This assumption is reasonable for the data. Additionally results are robust to wide range of values.
C. Prescriptions of Theoretical Models

Substituting the binary dynamic reward function of either winning or losing the tournament, for the continuous single period reward function of figure 2.1, results in substantial changes in the prescriptions of the two models as shown in figures 2.2 and 2.3.

While the Bellman of equation 2 is clearly a better fit for optimal tournament behavior, it is not clear that the dynamic model better represents players’ actual decisions. There are a couple of reasons to maintain this skepticism. First, this specific decision accounts for only five percent of the total decisions made by high-stakes gamblers. So despite the skills they have acquired over their careers and demonstrated by getting to this stage of the tournament, it is possible these gamblers have earned spots at the final table of a televised tournament, by maximizing other decisions. Second, players are allowed less than a minute to make their choice. This time constraint in conjunction with the extremely large statespace of the game (even in the final river round) may necessitate some heuristics rather than solving the Bellman equation. Maximizing the expected stack at the end of the hand appears a logical choice. Third this type of static maximization, albeit modified with rules of thumb to attempt to account for tournament poker’s credit constraint, is proposed by top players.

The earlier quote of Sklansky can easily be interpreted in a static expected utility framework as approximating risk aversion. Following this advice, tournament players
may behave as if maximizing their stack minus an opportunity cost of taking on more risk. A risk premium in the form of higher required expected value when bets are higher could compensate players for calling the current bet and risking forgoing future bets. However, this same behavior could also be interpreted in a prospect theory framework as loss aversion with the premium compensating players for the greater psychological costs of losing chips than the benefit of an equivalent gain. What this behavior can not be is risk aversion in the true neoclassical sense. Because a player must have all the chips to win the tournament (and the chips have no value outside of the tournament), the marginal benefit of any chip is identical to the next. Throughout the remainder of the paper, I do not attempt to differentiate between the two possible explanations for the same behavior. Conservative behavior may be the result of either loss aversion or a heuristic to act as if statically risk aversion in order to approximate dynamic optimization.

Another heuristic is offered by Doyle Brunson. Recognizing the difficulty of tournament play, Brunson proposed an ad hoc dynamic correction to the static optimization. "the strategy of tournament play differs from the strategy of ordinary play...I had been jeopardizing my chips on even-money situations, which can be a very good strategy in the early stage of an ordinary Poker game, but it is not good in a tournament, where you can't pull another thousand out of your pocket and buy more chips. In my new strategy, I tried to avoid playing big pots until the field had been narrowed substantially." In essence, Brunson is warning a lay poker audience about the
difference between optimization in a market with efficient credit markets and the credit
constrained environment of tournament poker. His solution however falls soundly
within static optimization.

One question is how closely these static heuristics approximate the dynamic
optimization of tournament poker. Figure 2.2 presents the indifference curves resulting
from optimizing several single period objective functions and is shown in contrast to
figure 2.3, the indifference curve of the dynamic bellman equation. The static
indifference curves of Figure 2.2 are presented both with the theoretical risk neutral
value and players’ revealed risk/loss aversion as calculated in section 2.3. To facilitate
comparisons between the two figures, figure 2.3 is calibrated to the empirical
distribution of state variables as also described in section 2.3. Both figures are
presented for games with four players left around the table.

First looking at the static optimization of figure 2.2, the Y axis is the risk neutral
expected value of calling (the sum of the top two branches of figure 2.1, weighted by
their expected probabilities). The X axis is the amount a player must Bet to stay in the
hand. To facilitate graphical interpretation both the risk neutral EV(call) and Bet are
presented as a fraction of a player’s Stack.\footnote{This normalization is natural within poker as both the costs and opportunities available to a player are defined in terms of the number of chips a player currently has.} The risk neutral static optimizer is indifferent between calling and folding only when the EV(call) = 0 along the X axis. Allowing for loss aversion or heuristics that approximate risk aversion changes the indifference curve into a monotonically increasing function of Bet.
The loss averse player requires a premium (measured as the distance above the X axis) which is a monotonically increasing function of Bet. The premiums required are measured in the graph with dotted lines at the observed mean Bet of 19% of a player’s chips. The size of this premium and whether the indifference curve is linear or convex depends on whether players’ rate of aversion is assumed to be constant or relative to stack or bet. A linear premium corresponds to neoclassical absolute risk aversion (CARA) whereas a nonlinear premium comes from behavior approximating relative risk aversion (CRRA). In figure 2.2, we can see that a player with relative aversion (RA) requires an increasing premium as investment costs risk, which is equivalent to either loss aversion or the typical specification of CRRA.\textsuperscript{18} A fourth indifference curve (RA\textsubscript{sub}) is presented in Figure 2.2. This curve assumes players believe their odds to be worse than the naïve monte carlo odds, E(prob) ≤ prob. Under this assumption the same observed behavior could be explained with a smaller premium.

For all models the indifference curves are continuous, monotonic and pass through the origin. With four players remaining in the game this can be considered a special case of Brunson’s recommendation to play all size pots with positive expected value once the field has been narrowed substantially.

In contrast to figure 2.2, figure 2.3 presents the indifference curve of a dynamic optimizer under similar realizations of the states. Graphing the indifference curve over

\textsuperscript{18} Bet/Stack is an inverse function of wealth
three dimensions – pot, chips and prob - is equivalent to taking a slice of the larger space where the remaining state variables, \textit{bet} and \textit{n}, are constant. This particular slice was chosen at the means of the observed data for the other two state variables and is representative of the larger function.\footnote{The non-monotonicity of figure 3 corresponds to the non-monotonicity of a binary choice dynamic policy function. The value function (ie the probability of winning the tournament) is reassuringly smooth and monotonically increasing (decreasing) in stack, prob, pot, and \textit{n} (bet).}

Comparing figure 2.3 to figure 2.2 shows some of the differences between the dynamic and static optimizations. Along the X axis, the indifference curve is no longer a monotonic function of a player’s chips. Instead the dynamically optimal choice roughly mirrors static risk neutrality when players hold between a third and three quarters of all the chips at the table. However at lower levels of wealth the dynamic model prescribes, as predicted by Sklansky, that even a risk-neutral dynamic optimizer will sometimes fold a hand with positive EV in order to preserve her chips for a later more favorable hand. There are also areas where risk neutral players will make bets with a negative EV because they don’t have enough chips to wait for better opportunities.

Preferences are monotonically increasing over both probabilities and pot for any given wealth level as would be expected, but the complexity of dynamically optimizing over different wealth levels doesn’t allow for an appropriate static simplification of the dynamic optimization over a range of \textit{Stack}. However, if a player or theorist were to oversimplify the prescription of the dynamic model with a heuristic that allows for decisions to be made in a static EV framework, the prescription could be construed as
risk aversion.\textsuperscript{20} For the purposes of this paper, I call this quasi-risk aversion. In this simplification, a player would behave as if she is slightly risk loving when the required bet is small relative to her stack: risk neutral over moderate size bets: and increasingly averse to bets representing a larger fraction of her wealth. This quasi-risk aversion raises the possibility that a player who is maximizing a dynamic value function may appear to be statically optimizing a risk averse (or even risk loving) utility function, as suggested by Sklansky.

However, even allowing for heuristics, as done in the models of figure 2.2, there are differences in the prescribed choices of the dynamic and static maximization. While Section 2.3 estimates players’ average behavior, Section 2.4 compares the outcomes of players’ decisions against both the theoretical models of this section and the following empirical models.

\textbf{2.4 Empirics}

\textbf{A. Description of Data}

The World Poker Tour (WPT) broadcasts a series of high-stakes Texas hold‘em tournaments every year. These broadcasts are later available for purchase on DVD. Each year’s tour consists of 12 tournaments where thousands of players pay between

\textsuperscript{20} This can be seen in figure 3 by looking at the general trend of the indifference curve. For a single value of prob the expected value that makes a player indifferent is generally increasing in bet/chips.
$10,000 and $25,000 to enter a tournament whose final payout ranges between $320,000 and $2 million. The lion’s share of the prize goes to the first place winner with lesser amounts going to the top eight finishers. With hundreds of participants in each tournament, players are initially randomly assigned to tables of eight. As players’ lose their stack of chips and are knocked out of the tournament, the tables are consolidated. After days of competition, a single table of six players remains. The DVDs consist of the action at this final table.

Broadcasting just the final table insures that our data come from skilled players. While anyone is allowed to enter the tournament, the final six players are the top of a true meritocracy, having knocked out hundreds of competitors in a zero-sum game. Interestingly, across the three years of broadcasts, a majority of the players at the final table come from a small cadre of professional high-stakes gamblers. The remaining players, while maintaining an outside profession, have mostly competed in high-stakes poker for years or decades.

This combination of characteristics of the WPT provides a unique opportunity to empirically test for optimization in an existing economic environment. I maintain that the observed decisions of these top high-stakes gamblers, competing in a zero-sum game, with tens of thousands in costs and millions in rewards will provide an upper bound on the optimality we can expect of decisions made in a dynamically complex, risky and uncertain environment.
To test players’ actual decisions and calibrate the models of Section 2.2, we
watched and coded three years of broadcasts from the WPT.\textsuperscript{21} Modeling the optimal
and observed decisions in the river round of betting, requires both the state variables
from the 124 choices and the expected distribution of these states in the entire
tournament. I combine the data from the 124 specific situations with the distribution of
variables from the entire 2,397 situations to calibrate the dynamic model.\textsuperscript{22}

In analyzing the data and following the advice of professional gamblers, I find
that the three state variables, Pot, Bet, and $E(Prob)$ and the parameter representing sunk
costs are functions of the number of players ($n$) left at the final table. These states and
the parameter are found to be monotonically decreasing in $n$. This corroborates poker
commentators’ reference to ‘the cost of poker’ going up as players are eliminated. It
also corresponds to the behavior observed in other natural monopolies, where both
investment costs and profits increase as the number of companies decrease (Sutton,

In the first hand of the final table, $n=6$, with players stochastically eliminated
until the final choices are made with only two players left at the table. The parameter $ave$
is set at the mean for each of the five values of $n$. The stochastic distributions of the
second, third and fourth state variables are modeled discretely using their histograms for
each value of $n$. The fifth and final state, Stack, is independent of $n$, taking on values
ranging from all the chips at the table to a single chip.

\textsuperscript{21} Thanks to Mike Brady, Yun Jae Hwang, Karl Meussen and Brian Roe for help in the coding effort.
\textsuperscript{22} Histograms of the state variables and sunk costs for both the 124 river decisions and 2,397 situations
that make up all 4 rounds of betting are available upon request.
Matching the 124 observed realizations of states to the appropriate combination of stochastic distributions is trivial for all states except $E(Prob)$. Because a players’ expectation is never observed, I substitute this variable with a player’s monte carlo odds of having a better hand than a randomly selected opponent given their two hole cards and the five community cards.

As shown in the remainder of this paper, these monte carlo odds are only a fair approximation of players’ true expectation of having the highest ranked cards. In the following regressions, these monte carlo odds are used as a naïve baseline with a player allowed to subjectively update her expectation using a single parameter power function - $E(prob) = prob^a$. This formulation is not found to increase the ability to predict players’ actual decisions.

According to comments from these players, their true expectation is much more situation specific then allowed for by either the naïve monte carlo odds or the power function. However far from being a detriment, sticking to the naïve monte carlo odds allows Section IV to test and reject the efficacy of players true expectation and subsequent decisions to outperform any of the models in terms of frequency or amount won. While the monte carlo expectation of winning is only a fair representation of players’ true expectation, it turns out their true expectation contains more noise then information.
B. Estimates and Inference

Using data from the WPT, I estimate parameterizations of section 2.2’s objective functions to match the actual decisions of high-stakes gamblers. The regressions provide estimates of the representative player’s risk/loss aversion, her expected odds of winning and her use of rules of thumb that bring dynamic considerations into a static framework. Given these estimates, I test and find 1) conservative behavior approximating risk/loss aversion, 2) players’ expectation of having the best cards uses the naïve monte carlo odds but also contains outside information or noise not included in the theoretical models.

Based on the models of Section 2.2, I estimate the static objective function that when maximized, best matches the observed decisions of professional gamblers. These static models assume players are trying to maximize a function of expected stack. While the single period theoretical model simply maximizes expected stack, the comments of professional gamblers suggested two modifications to this model. Allowing for loss/ quasi-risk aversion and an intercept specific to the number of players, the static regressions better capture observed behavior than risk-neutral theory.

Estimating a single measure of aversion for players in different chip positions across different tournaments requires a normalization of the costs and benefits they face. This is because under quasi-risk aversion the marginal value of a chip is different depending on the number of chips a player has in her stack, while under loss aversion the size of the potential gain and loss relative to a player’s current position is what
matters. Luckily the rules of poker provide an easy normalization across players of the costs and benefits. Just as in the static graphs of section 2.2, dividing all costs, bets, and all potential benefits, pot, by a player’s stack, provides a measure of cost and benefit that is generalizable across players. This makes sense within tournament poker for a couple of reasons. First chips have no value outside of the tournament, avoiding issues of outside opportunities for players with different stacks. Second all bets and subsequently the pot are presented to players relative to their stack. A player is only allowed to bet the chips she has, and she is always allowed to bet all her chips. Similarly the amount a player can win is a function of the amount she bets (as shown in Fig 2.1). In addition to providing a way to measure professional gamblers aversion to large bets, measuring costs and benefits as a fraction of stack is how players always refer to their decisions. Commentators, advice books and players themselves refer to the fraction of their chips required to call and the percentage increase in their chips if they win.

While the marginal cost and benefit of a single chip are equivalent for a static risk neutral optimizer – risk aversion is equivalent to decreasing marginal utility of chips. The typical neoclassical explanation for this nonlinearity doesn’t hold for tournament poker, as chips have no value outside of the tournament and as shown in section 2.2 chips enter into players’ reward function linearly. Instead the decreasing marginal value of chips may better be understood as either loss aversion as presented by Prospect Theory (Kahneman & Tversky 1979) or quasi-risk aversion as presented by
Sklansky. Under loss aversion, losses and disadvantages are weighted more than gains and advantages for psychological reasons. In the framework of the choice to fold or call quasi-risk/loss aversion can be captured by allowing the cost \( Bet \) and benefit \( Pot + Bet \) to enter into players objective functions differently. Simplifying and parameterizing the static objective function (1), per the binary decision equation (3), allows estimates of the level and type of the representative player’s aversion as well as her use of subjective weighting of probabilities and heuristics to bring dynamic considerations into a static optimization.

\[
d = 1 \left( \text{prob}^a \geq \frac{\hat{\theta}_2}{\hat{\theta}_1} \frac{bet^\gamma}{bet + pot} - \tilde{e} \right)
\]

Under absolute aversion the ratio of the disutility of costs (\( \hat{\theta}_2 \)) and utility of benefits (\( \hat{\theta}_1 \)) is constant throughout the range of possible costs, resulting in an estimate of \( \gamma = 1 \). In this case the premium required by a player to bet is simply \( \frac{\hat{\theta}_2}{\hat{\theta}_1} \times bet \). With relative aversion the premium is generally assumed to be increasing in the size of the bet. The differences in the premiums required under absolute and relative aversion were presented graphically in figure 2.2 and are captured in the econometric framework by
the estimate of $\hat{\gamma}$. Not placing restrictions on the values $\hat{\theta}_1, \hat{\theta}_2$ and $\hat{\gamma}$ can take, allows for the possibility that players’ choices may best be represented by uncommon preferences such as decreasing relative aversion or even risk loving behavior. While it may be difficult to imagine risk loving behavior in professional decision makers, many would consider entering a zero sum game with large costs and huge payouts the definition of risk loving. I leave it to empirics to resolve this issue.

Table 2.1 presents the logistic estimates of three versions of equation 3, assuming type II extreme value errors.

<table>
<thead>
<tr>
<th>Variables</th>
<th>RA(sub)</th>
<th>RA</th>
<th>AA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.70 (.62)</td>
<td>.59 (.71)</td>
<td>.67 (.45)</td>
</tr>
<tr>
<td>5 or 6 players</td>
<td>-.98 (.57)</td>
<td>-1.00 (.55)</td>
<td>-.99 (.55)</td>
</tr>
<tr>
<td>3 or 4 players</td>
<td>-.07 (.57)</td>
<td>-.06 (.56)</td>
<td>-.04 (.55)</td>
</tr>
<tr>
<td>$\theta_1$ Pot + Bet</td>
<td>4.97 (1.1)*</td>
<td>4.03 (.89)*</td>
<td>4.03 (.88)*</td>
</tr>
<tr>
<td>$\theta_2$ Bet</td>
<td>-8.13 (3.7)*</td>
<td>-9.21 (3.1)*</td>
<td>-9.11 (2.7)*</td>
</tr>
<tr>
<td>$\gamma$ Relative Aversion</td>
<td>1.24 (.66)</td>
<td>1.05 (.42)</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha$ Subjective Prob</td>
<td>3.07 (1.3)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N=124</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln Likelihood</td>
<td>-66.65</td>
<td>-69.93</td>
<td>-69.95</td>
</tr>
</tbody>
</table>

Notes: "*" denotes coefficient estimate is significantly different than theoretical value at the p<.05 level
First finding – Conservative Choices:

Within the framework of the game and a static optimization, professional gamblers require a significant premium to bet. Hypothesis 1 sets $\hat{\theta}_1 = \hat{\theta}_2 = \hat{\gamma} = 1$ in order to test the null of risk-neutrality. A LR test is rejected with greater than 95% confidence for all three models presented in Table 2.1. However, the specific form of loss/quasi-risk aversion is not identified\(^{23}\). Comparing the three models, we find a relatively small difference in their statistical properties corresponding to a noticeable difference in the premiums required to call (figure 2.2’s indifference curves). While we can reject players behaving as if risk neutral or risk loving, the standard errors don’t allow identification of the exact shape of risk/loss aversion. This finding is formalized in Hypothesis 2, where Wald & LR tests fail to reject statistical differences between the AA and RA models. Whether these decisions are due to loss aversion or a response to dynamic credit constraints\(^{24}\), the next section shows that they result in significantly more conservative outcomes than either static or dynamic optimization.

Given the variable returns and investment costs of high stakes poker, the premiums required by professional gamblers are naturally measured over both dimensions, as in figure 2.2. The indifference curves of figure 2.2, built from the

---

\(^{23}\) Ln likelihood of 3 restricted models: -82.38, -83.25, -83.25

\(^{24}\) As presented in Section II, the required risk premium could either be due to loss aversion or heuristics used by players (quasi-risk aversion) to approximate dynamic optimization within a static framework. Under either quasi-risk aversion or dynamical optimization, the curvature of the estimated function may represent the curvature of the dynamic value function rather than the static utility function. In this case the Arrow-Pratt measure of risk aversion would be capturing $-\frac{V'(w)}{V''(w)}$ in place of $-\frac{U'(w)}{U''(w)}$.  

69
regressions of table 2.1, contrast sharply to the revealed preferences of amateur gamblers. The professional gambler’s objective function (calculated via duality’s: $\frac{\partial y}{\partial x} = -\frac{\partial x}{\partial U(x,y)}$, is concave over chips. This concavity corresponding to quasi-risk or loss aversion contrasts the convex utility functions found in racetrack betters (Jullien & Salanie 2000, Ali 1977). Definitively testing whether this is due to differences in horse racing and poker or amateurs and professionals would require data over professional horse betting or amateur poker. While the similarities between racetrack betting (parimutuel) and poker (zero-sum) are strong, comparing the decisions of amateurs and professionals is not of primary interest to this paper (see Haigh & List 2005 for evidence of greater loss aversion in professionals than amateurs). Instead my interest is in the behavior of professionals in their own right. The indifference curves calibrated from the regressions suggest that professional gamblers are willing to invest just under twenty percent of their chips for between 1.2 and 1.3 expected rate of return, but require an expected return of close to two before risking half their chips. In other words, professional gamblers- a group that many assume to be risk loving - are behaving as if loss averse. Section 2.4 expands the analysis in order to test gamblers’ decisions against dynamic optimization in addition to static stack maximization. The results confirm and expand the findings that professional poker players are overly conservative.
Second finding – Signal trading

\[ E(\text{prob}) = f(\text{prob, outside information}) \neq \text{prob}_{\text{monte carlo}}. \]

In hypothesis 3, allowing for subjective probabilities fails to substantially improve the ability to predict players’ behavior (fail to reject \( \hat{\alpha} = 1 \)). Despite its inability to significantly improve estimates of when players will call, a subjective probability model is included for a couple of reasons. First, comparing models one and two (figure 2.2, table 2.1) shows both the difference and the significance of the required risk premium when the subjective odds of winning are systematically lower than the naïve probabilities.

Second, the low t-value of \( \hat{\alpha} \) doesn’t necessarily support the hypothesis that players’ expectations equal the naïve monte carlo probabilities. The formation of this expectation is the only uncertainty in players’ final decision optimization and therefore a likely source of differences in observed and estimated behavior. While the inclusion of naïve probabilities drastically improves the ability to estimate professional gamblers’ decisions, the inability of all three models to perfectly predict observed choices (seen in the log likelihood) points to a difference in players’ subjective probabilities of success.

A possible explanation is that the chosen specification poorly approximates players’ true subjective probabilities. Despite being a typical choice (Jullien & Salanie 2000) and having intuitive appeal to this specific poker decision\(^{25} \), the power function

---

\(^{25}\) Players often know when they have the best hand (the ‘nuts’) without having to see their opponent’s cards in which case \( E(\text{prob}) = \text{prob} \) when \( \text{prob} = 1 \) as is the case with the power function. Also players are no longer facing a randomly selected opponent as is assumed in the Monte Carlo odds. The power function allows for all subjective probabilities (especially those with the highest variance \( \approx .5 \)) to be
may not capture players’ true updating of the naïve odds. However, replacing the simple function with up to a fourth order polynomial expansion (which allows for expectations more in line with the predictions of prospect theory) also fails to increase accuracy.

An alternative explanation is that players are using signals, in addition to the naïve odds, in forming subjective probabilities of having the best cards. This explanation is supported both by players’ explanations of their behavior and the importance of the error terms.

Despite the clear significance of the variables included in the regression and the conclusions already drawn from the ability of the model to fit the data (loss aversion), I also infer something from the inability of the model to fit the data perfectly. By looking only at the significance of the model as a whole and the variables of economic importance specifically, we may miss the importance of the error term. The error terms explain why the log likelihoods, show a clear divergence of players’ behavior from that predicted by the models.

The error term captures the effect of either model misspecification or variable misspecification. While the model includes all the variables of economic significance for static choices, the error term may capture players’ attention to non-economic

---

26 In this usage variable misspecification does not refer to measurement error. The variables included in the analysis Bet, Prob, Pot, Stack are unambiguously presented in the televised poker tournaments. Unlike other datasets, coding decisions are not required. Instead variable misspecification refers to players’ subjective probs as E(prob) is a key unobservable variable in the optimization.
variables, differences in their expectations of having the best cards, or the possibility that they are dynamic not static optimizers. The clear expositions of EV in poker advice books and the explanations of actual decisions by professional players themselves make the first explanation unlikely. Similarly descriptions of static optimization in the same sources dissuade me from the third explanation. None the less, the dynamic theoretical model of section 2.2, allows for testing of this hypothesis. The second explanation that players are forming subjective expectations, significantly different from the monte carlo odds, seems the most likely.

The previous hypothesis tests, rejecting players’ systematic updating of naïve odds, do not allow for players’ use of additional signals in forming their expectations. These situation specific expectations are precisely what players claim to use. After offering brief explanations of naïve probabilities and their role in maximizing expected value, players’ interviews and books spend an inordinate amount of time on how they outperform/update these naïve probabilities. The proposed rules involve identifying opponents’ physiological tells and multiperiod analysis of their bets.

Unlike the variables and decision rules used in this paper, both the factors involved in this type of subjective updating and how they should be weighted are hotly contested. Instead of trying to test for the specific type of signals used, I take advantage of a unique characteristic of the data to test the efficacy of using these signals. Given players’ decisions are different from those proposed by the theoretical and empirical models, Section IV tests if the divergence helps or hurts. Finding that the signals used
by players contain more noise than information provides another reason to limit our modeling of these situation specific expectations.

**Ancillary findings – overcoming non-stochastic behavioral shortcomings & dynamic considerations**

Additional implications of professionals’ decisions are found in the regressions’ constant terms. Estimates of a single intercept are insignificant and have an extremely small point estimate\(^{27}\), indicating professional gamblers have no predisposition to betting. Despite having bet in the three previous rounds with the same hole cards – professionals avoid the sunk cost fallacy (throwing good money after bad). Replacing this constant with measures of sunk cost in the hand does not change the result, even though the sum of previous bets account for on average fifteen percent of players’ chips. The ability to walk away from large, self-inflicted costs is reported by both the economic and psychology literature to be the most difficult type of sunk cost fallacy to avoid (Staw & Ross 1989). While in the lab, subjects are found to fall prey to sunk cost (Cachon & Camerer 1996) and in field experiments average individuals are found to do the same (Arkes & Blumer 1985) – the size of previous bets have no determination on professional gamblers’ decision to fold or call. Player’s ability to do just this is highly praised by poker commentators and may partially account for the fact that a handful of professionals are able to make a living in a zero-sum game.

---

\(^{27}\) All three models: constant \(\approx -0.1\) with standard deviation \(\approx 0.5\)
As shown in Table 2.1, replacing the single intercept term with fixed effects for the number of players around the table does improve the models’ ability to explain gamblers decisions. Players are systematically less likely to call when more players (5-6) remain at the table when compared to their choices when two, three or four players remain. These estimates of intercepts specific to the number of opponents points to an attempt to approximate dynamic optimization within a static framework as proposed by Brunson.

Finding professional gamblers to avoid sunk cost fallacy but fall prey to overly conservative behavior may appear contradictory. Much like the literature, we find professionals avoiding one behavioral anomaly (Kagel & Levin 1986, List 2004) while simultaneously falling prey to another (Hogarth & Kunreuther 1989, Locke & Mann 2000, Haigh & List 2005, Romer 2006). However these findings allow for an explanation offered by research in psychology (Langer 1975) and beginning to be explored in economics. The ability of competitive economic environments to drive professional behavior to the neoclassical prediction, through either learning or natural selection, may depend on the risk involved in the decision and how consistently incorrect decisions are punished.

In Hold’em, players are required to make sunk investments in every hand in the form of antes and blinds. These bets must be placed before a player has a chance to look at her cards. Additionally the high variance of the probability of winning the
hand\textsuperscript{28} necessitates that any successful player place some bets she will later be forced to walk away from. While these circumstances aren’t quite as good as those of a risk free competition in correcting behavioral anomalies, they are a very close second.

Contrasted with this opportunity to overcome sunk cost is tournament poker’s limited opportunity to learn risk-neutral optimization. Both the dynamic nature of the tournament and the risk involved in the final decision create difficulties in players’ optimization. Recall that we are only looking at the decisions of players who have made it to the final table. These players have made avail of some strategy to eliminate hundreds of opponents in earlier rounds. While this paper models the dynamics of tournament play at the final table, I don’t attempt to prescribe behavior earlier in the tournament. One major reason for this is, unlike players at the final table, all the players who enter the tournament can not reasonably be assumed homogenous. Anyone with $10,000 is allowed to enter the tournament regardless of skill or experience. Many of these entrants are competing merely for fun with little chance of winning. Advancing to the final table when competing against hundreds of heterogeneous players may well require choices resembling risk aversion.\textsuperscript{29} This raises the possibility that players are applying the same strategies that got them to the final table to the final table competition. In this case, both the opportunity for learning risk-neutrality and competition selecting for risk-neutrality are limited. Additionally the risk of the final decision at the final table works against natural selection. In this case players may often

\textsuperscript{28} See appendix A  
\textsuperscript{29} This is the essence of both Sklansky & Brunson’s books
be rewarded for making overly conservative decisions (seeing that they would have lost in the showdown) and punished for making the right decision (losing chips on a positive EV bet). In this way we see how players may simultaneously avoid sunk cost, while adopting an overly conservative style of play (learning the wrong lessons or applying the right lessons to the wrong situation).

In addition to confirming the suboptimality of gamblers’ observed decisions and rejecting the ability of heuristics to approximate dynamic optimization, the next section tests the optimality of the second major empirical finding of this section. Players’ complex situation specific expectations are shown to contribute nothing over simple statistics to either their number of correct decisions or the resulting change in stack. The implication that players’ updating of probabilities contains more noise than information is explored.

### 2.5 Outcomes

One of the factors leading to the popularity of televised high stakes poker, is the use of lipstick cameras. These miniature cameras built directly into the poker table allow all players’ cards to be broadcast, while around the table players only see their own cards. The goal of this broadcast information is to make an extremely stochastic and uncertain game more accessible to the casual viewer. It also has the unintended benefit of both
making the empirical tests of the previous section possible (*because we can see players’ hole cards*) and allowing an important extension to the analysis of the first four sections (*because we can also see opponents’ hole cards*).

In the first 3 rounds of betting, a player who knew her opponents’ cards would have an extreme advantage. Knowing all the (hole) cards at the table would remove the uncertainty of the decision. However, in the first rounds of betting the player would still face risk as community cards remain to be dealt.\(^{30}\) Even removing the uncertainty of opponents’ cards and precisely defining the full information probabilities of winning (as is done for viewers of the WPT), the high variance of Hold’em maintains interest and suspense. This is true not just for viewers but for players. Occasionally players see their opponents’ cards before the showdown when further bets can not be matched because all but one of the remaining players have gone “all in”. In these circumstances, players’ relief (or frustration) is evident at learning with certainty the risk they face.

However this is not the case in the final river round of betting of interest to this study. Like most economic decisions in the real world, a player must decide to call or fold facing the uncertainty of both her opponent’s (hole) cards and thus her true probability of winning in the showdown. While the theoretical models of section 2.2 bypassed this uncertainty by assuming opponents’ cards are randomly selected, the empirical models of section 2.3 found players’ decisions to diverge from those of either a risk neutral or conservative model using this naïve expectation of winning the hand.\(^{31}\)

---

\(^{30}\) See Appendix A for game description

\(^{31}\) Although the conservative model was a significantly better explainer of players’ decisions
This section uses the full information odds of winning the hand (known only to the viewer) to compare players’ observed outcomes against the outcomes attainable under the previous sections’ theoretical and empirical models.

In the final river round of betting, knowing all the cards at the table would make the decision to fold/call trivial. This is because all the cards have been dealt and the player’s true probability of winning ($\text{prob}$) is either 1 or 0. If $\text{prob} = 0$ then the player would fold and if $\text{prob}=1$ the player would call.

But in high-stakes poker, the players don’t have this information and we are not interested in what they would do if they did. It would be like predicting a stock analyst’s recommendation to buy or sell assuming he holds a company’s next quarter profits in his hands. Instead we are interested in how well professionals form their expectations of success and the optimality of their observed choices. Testing the optimality of players’ choices against the generalized decision rules of sections 2.2 & 2.3, this section is analogous to tests of stock analysts’ earnings. Like those tests observed earnings are compared to the risk adjusted profit of decision rules that use only a subset of the information available to the analyst. I find, like studies of behavioral finance, that even professionals act on noise as if it were information and that their decisions are overly conservative.

Unfortunately noise not only affects decision makers but the econometrician as well. While we are ultimately interested in professional gamblers’ overall success rate, aggregating different classes of decisions can obscure information. For this reason,
binary econometric models use percentage concordant tables to show models’
divergence from both of the observed choices. Just as in percentage concordant tables, I
compare observed and modeled decisions for two classes of final round bets.
Successful choices can come from only one of two scenarios – call when a player has
the best hand, fold when she doesn’t.

Using the full information available to the viewer, Figures 2.4 and 2.5 divide the
124 situations into when a player should call and when she should fold. The
corresponding tests reject players’ outcomes coming from static or dynamic
optimization. Instead both in terms of average amount won and lost, professional
gamblers’ decisions correspond to significant loss aversion. Players’ complex rules for
‘reading’ opponents shows no improvement over decision rules that assume opponents’
cards are random.

Both the average probability of correctly calling and the resulting average
percent increase in Stack are presented in Figure 2.4. The outcomes are divided into the
three empirical models of Section 2.3 and three theoretical models. The fourth and
sixth models correspond to the static and dynamic optimizations of Section 2.2. The
fifth model slightly modifies the static model of section 2.2 by replacing the naïve odds
of success with the estimates of the power function from section 2.3. For both
probability and amount won, the two panels’ range has been limited to the maximum (1,
18%) and minimum (0,0%) possible. The outcomes of a player who was always right,
choosing to call when the lipstick camera shows she will win, would appear on the
upper bound of both panels resulting in 100% correctly calling and increasing her Stack by on average 18% in the river round of betting. In contrast, players’ actual decisions are represented by the horizontal line crossing all six models. The distance between the players’ observed outcomes (both prob of correct decisions and amount won) and the diamond of each model is the difference in outcomes. The diamonds appear at the model means with bars indicating plus and minus one standard deviation of the mean difference between the model and players’ decisions. The standard deviation of the difference between players’ observed outcomes and theory emphasizes questions of how closely observed decisions approximate optimization.

Doubling the length of the bars provides a measure of which models’ outcomes significantly outperform the average decisions of players.\(^{32}\) While the intuition of this test turns out to be correct, the appropriate statistical tests of the difference in the models from players’ decisions are slightly more complex.

Testing whether players make significantly fewer or more correct decisions than the models requires two procedures - one for the empirical models and one for the theoretical. There is nothing stochastic about the theoretical decisions or what we can observe of players’ decisions. Theory and the players either fold or call with absolute certainty. However, how often these prescribed or observed decisions match the correct choice (as known by the viewer) is stochastic. As I only differentiate between the 124 final round bets based on whether the lipstick camera shows the player will win or lose,

\[\text{Graphical representation of a t-test with 96% confidence}\]
the distribution of players’ and theory’s correct decisions to call is best modeled as
Bernoulli. The likelihood that players’ observed correct decisions came from the same
Bernoulli distribution as each of the three theoretical models is tested using the
observed means and variances of the models and observed behavior. A Pearson $\chi^2$ test
rejects with greater than 99% confidence that players’ observed decisions to correctly
call are coming from any of the three theoretical models. Professional gamblers play
significantly fewer winning hands, consequently winning significantly less chips, than
either a static hand or dynamic tournament maximizer. This result over the
suboptimality of outcomes corroborates and expands the findings of section 2.3.

Interestingly in players’ comments, both static EV and differences in static and
dynamic maximization are clearly explained. However the heuristics players propose
and are found to use to deal with dynamics are counter productive. The outcomes of
simple risk neutral static maximization better approximate dynamic optimization than
players’ actual decisions.

Testing whether players’ outcomes are in line with loss aversion requires a slight
change to the Pearson $\chi^2$ procedure. While the Bernoulli distribution remains the best
fit of players’ correct decisions, the empirical models estimate a probability of calling
for each decision. These probability estimates of the logit models are distributed as a
poisson trial. Even with the lower variances of a poisson trial, Pearson $\chi^2$ tests fail to
reject that the observed outcomes came from any of the three risk averse models.$^{33}$

$^{33}$ For a fuller exposition of the poisson trial specification see Appendix B
The second panel in Figure 2.4, shows the amount won from calling as well as the differences from observed behavior. Wilcoxon Matched-Pairs Sign and Signed-Rank tests are performed comparing the chip changes of the players with those resulting from the model prescriptions. The amounts actually won by the players (0 in the case of incorrectly folding and \( \text{Pot+Bet} \) in the case of correctly calling) are found to be significantly less than the amounts they would have won following any of the theoretical models. Again the Wilcoxon nonparametric tests fail to reject that the amount won by players differs from the empirical risk-averse models.

Over both metrics shown in figure 2.4, players fail to outperform the loss averse empirical models of section 2.3. The implication for players’ ability to read their opponent is not good. Significant differences in the choices of professional gamblers and the prescriptions of loss averse models based on players’ mean behavior do not correspond to differences in outcomes. Ignoring complex situation specific rules for calculating the odds of having the best hand does not hurt the models. As will be confirmed in the next figure, any signal players are using in forming their expectations of success, beyond simple statistics, is more noise than information.

As in percentage concordant tables, a model may match observed behavior well along one dimension (correct calls) but poorly along the other (correct folds). Or two models may have similar overall accuracy corresponding to vast differences in accuracy along the two separate dimensions. However, like Figure 2.4, Figure 2.5 shows a significant difference in players’ observed outcomes and the outcome of following any
of the theoretical models. In this case, I test and reject that both the number of players’ correct decisions to fold and the amount lost when they should have folded were generated by risk-neutral theory. As shown in the first and second panels of Figure 5, finding theory to yield more and larger loses than players’ observed decisions is reassuring given that I also found theory to result in more and larger wins. As expected, risk-neutrality results in both more won and lost than either loss aversion in the three empirical models or quasi-risk aversion of observed behavior.

Jointly figures 2.4 and 2.5 reveal an unexpected finding in this specific decision – the mean outcome of professional gamblers who have paid thousands of dollars to enter a zero-sum game, match those of loss aversion. While finding loss aversion in any decision would be striking, it is startling in the river round of betting where the realizations of costs, benefits and naïve probabilities result in expected value being positive more than 90% of the time. Despite this vast improvement over mean expected value, players are overly conservative calling only half the time. Pearson $\chi^2$ and Wilcoxon tests confirm players’ expected losses are in line with loss averse decision rules.

Figures 2.4 and 2.5 show the risk averse models match players’ mean outcomes extremely well. But the the variance between the empirical models and players’ actual decisions doesn’t allow us to accept with high confidence\(^{34}\) that they are the same. Of the 86 observations when players should have folded, the empirical estimates of

\(^{34}\)The type I error, or $\beta$ is 35% meaning the power of the test remains small due to the variance in observed behavior and the empirical models. This reinforces the previous findings of the log likelihood and % concordant tables of section III.B
players’ decisions under risk aversion only match 65% of their actual decisions. As we have rejected both risk neutrality and dynamic optimization, this leaves players’ formation of the expected probabilities of winning as the source of discord.

Players’ subjective probabilities are neither systematically higher nor lower than the naïve monte carlo odds as shown by the failure to reject the naïve odds models of section 2.2. Players are just as likely to call when the risk averse econometric models predict folding as they are to fold when the models prescribe calling. Supporting this finding are the decision rules described by professional poker players themselves. Gamblers report the naïve monte carlo odds playing an important role in their subjective probabilities, however, they also report using complex situation specific updating of these odds. It is this updating of odds that leads to differences in observed decisions and empirical models.

The results presented in FIG 2.4 and FIG 2.5 and the Wilcoxon and Poisson hypothesis tests provide us with information as to whether the naïve monte carlo odds or players’ true expectations are more accurate. Players’ outcomes using complex situation specific expected probabilities of winning are statistically no different from the risk-averse models. In fact these simple decision rules that imitate risk or loss aversion by requiring a premium increasing in the size of the bet result in both less risk and slightly greater gains than players’ observed decisions.

Having found high-stakes gamblers’ choices to be more conservative than both statically and dynamic optimal and rejecting the ability of their heuristics to either
improve their ability to deal with credit constraints or estimate the odds of having the best hand, Figure 2.6 summarizes players’ outcomes.

The overall results of all 124 observations averaging losses and gains are presented in Figure 2.6. While on average following any of the theoretical models would double the expected winnings of gamblers in the final bet of the hand, the increased variance resulting from summing up loses and gains makes this difference statistically insignificant at the typical confidence levels. This masking of significant differences both in the domain of loses and wins may help explain the divergence of players’ observed decisions from theory.

Research shows that making optimal decisions in the face inconsistent feedback is incredibly difficult (Langer 1975). In the overly conservative choices of high-stakes poker players, as in similarly suboptimal decisions of professional football coaches (Romer 2006) or options traders (Height & List 2005), decision makers receive positive feedback for suboptimal decisions. The risk inherent in these decisions means that players often win a bet they shouldn’t have placed or lose a bet with positive expected value. Given the body of evidence in both neoclassical economics and prospect theory that individuals weight the marginal cost of a loss more than the marginal benefit of an equivalent gain, it is easy to see how even professionals may take away the wrong lessons from their decades of experience to make overly conservative decisions. The opportunity that both static and dynamic optimization provide to double the percentage
of chips won over observed decisions comes at the cost of almost doubling expected loses.

Additionally the fact that this specific decision accounts for only 5% of professional poker players’ choices may limit the ability of competition to force behavior to the neoclassical maximum. The opportunity cost of a player’s conservative choices in the final river round of betting only averages out to 4% of her stack. While this is a noticeable disadvantage over the three years of data analyzed, it is easy to see that an optimizer could still easily lose the tournament due to poker’s stochastic nature despite making the correct choices in this specific decision.

Finally the finding that players use complex situation specific signals in addition to simple statistics in forming their expectations of having the best hand, is also in line with cognitive findings (Kahneman & Lovallo 1993). Poker players’ divergence from even a simple loss averse maximization is in keeping with decision makers’ tendency to view situations as unique and therefore supplement or in many cases replace their statistical probabilities of success with ‘inside’ views of their likely outcomes. The fact that multiperiod analysis and psychological tells don’t improve players’ performance at the final table of the World Poker Tour, doesn’t prevent players from treating this noise as if it were information.

These ready explanations for overly conservative behavior and confounding noise for information negate neither the significance nor implications of the findings.
2.6 Conclusion

Central to expected utility theory are competitors’ abilities to form rational expectations of the future and to use these expectations in maximizing profit. While the efficient functioning of markets is predicated on these two skills, empirical evidence has called these assumptions into question. Behavior observed in the laboratory systematically diverges toward overly conservative choices and overly confident expectations. While these findings have led some to propose alternatives to expected utility theory (Savage 1958, Kahneman & Tversky 1979), others (Friedman 1953, Haigh & List 2005) point out that these findings come primarily from laboratory studies of inexperienced subjects competing for small stakes.

Providing several insights into the ability of competition, large incentives and learning to bring out optimization, I analyze a specific class of decisions by professional poker players. Skilled gamblers competing in high-stakes tournaments are found to overcome behavioral anomalies (sunk cost fallacy) with clear and consistent repercussions. However when facing uncertainty and dynamic risk, players make significantly more conservative choices than either statically or dynamically optimal. This behavior is consistent with either loss aversion or the use of intuitive but counterproductive heuristics. Additionally complex expectations used by players to
overcome the uncertainty of who has the best cards fail to improve outcomes over simple statistics.

These findings from one of the highest-stake competitions in existence are in line with both laboratory findings and persistent market anomalies (Mehra & Prescott 1985, Roll 1986). While some studies have found behavioral anomalies attenuated by competition and learning, the persistence of human foibles points to a limit of economic rationality. The risk, uncertainty and dynamics inherent to tournament poker, and so much economic decision making, limit players’ ability to optimize simple objective functions. Despite being chosen as ideal economic agents for normative explanations of rational expectations and expected utility theory, positively poker players diverge significantly from neoclassical predictions.
References


Locke, Peter, and Steven Mann, 2000, “Do professional traders exhibit loss realization aversion?” working paper, Texas Christian University.


35 A virtually identical story about experts’ miscalculation of a specific type of probability is found in Brunson 1978, Savage 1954, and Bernouilli 1700. Apparently this is a human trait.
Figure 2.2 Static indifference curves of loss averse and risk neutral poker player
Figure 2.3 Dynamic indifference curve of loss averse and risk neutral poker player holding bet =2 at the mean observed value for games with 4 players remaining
Figure 2.4 Graphical & Statistical Comparisons of Outcomes when player has best hand

Notes: 1. The first panel shows the probability of correctly calling for the empirical models (first 3 bars), the theoretical models (last 3 bars), and the observed behavior (horizontal line). The second panel shows the resulting increase in gamblers’ chips. A diamond marks the mean. The vertical lines represent a standard deviation of the difference between the models and observed behavior.
2. The significance of the difference between players’ decisions and the models is measured with the Pearson $\chi^2$ test for probabilities of success and with the Wilcoxon Signed-Rank test for the amount won.
3. “*” denotes gamblers’ actual decisions significantly underperform theory at the $p < .05$ level. 4. N=38.
Figure 2.5 Graphical & Statistical Comparisons of Outcomes when player has worst hand

Notes: 1. The first panel shows the probability of correctly folding while the second panel shows the amount lost from incorrectly calling. 2. The $\chi^2$ statistics are presented for the first panel. The second panel includes the nonparametric S statistic. 3. "*" denotes dynamic and static economic theory prescribes choices significantly different from observed decisions at the $p < .05$ level. 4. N=86
Figure 2.6 Graphical & Statistical Comparisons of Outcomes over all hands

Notes: 1. Figure 6 summarizes the results of calling and folding presented in the previous two figures. 2. The maximum and minimum probabilities of success are bounded by the first panel. The upper and lower boundaries of the second panel represent the average percentage of her chips a player could win or lose. 4. N=124
Figure 2.7 Value function of average cards with six players
Figure 2.8 Value function of average cards with four players
Figure 2.9 Value function of average cards with two players
Texas Hold’em is a variant of poker, which is currently popular for high-stakes televised tournaments. One of the main reasons for this popularity is the extreme volatility of players’ odds of winning within a single game (or hand) due to the sequential revelation of new information.

A ‘hand’ of Texas Hold’em consists of four stages, each followed by a round of betting. A hand begins with each player receiving two private cards (hole cards). The next stage is called the flop, where three community cards (shared by all players) are dealt. One more community card is dealt in both the third stage (turn) and the fourth stage (river). Finally, after all four stages and four corresponding rounds of betting, all players remaining in the hand reveal their hole cards in the showdown.
The player with the best five-card poker hand formed from their two hole cards and the five community cards wins the sum of all bets (*pot*).
APPENDIX B
POISSON TRIAL

Following Nedelman & Wallenius (1986) the conditional expectations of success for the models can be defined by $E[S|E(\cdot), u(\cdot), X]$:

$$E[S \mid r_1, r_2, \ldots, r_n] = n\bar{r}, \quad (5)$$

$$\text{var}[S \mid r_1, r_2, \ldots, r_n] = n\bar{r}(1 - \bar{r}) - n\sigma_r^2, \quad (6)$$

where $\bar{r} = \frac{\sum_{i=1}^{n} r_i}{n}$ and $n\sigma_r^2 = \sum (r_i - \bar{r})^2$.

The differences between models come from the calculation of $r_i$. $r_i$ is the expected probability of making the right choice given either the maximization of the subjective expected utility function $(\hat{y}_i)$ or just the subjective expectation $(E(\cdot))$.\(^{36}\) In the first case:

$$r_i \mid \hat{y}_i = p_i^{\text{full}} \cdot \hat{y}_i + (1 - p_i^{\text{full}}) \cdot (1 - \hat{y}_i) \quad (7)$$

\(^{36}\hat{y}_i \equiv \text{expected probability of calling from the binary choice logistic model}\)
LIST OF REFERENCES

Chapter 1


Chapter 2


Locke, Peter, and Steven Mann, 2000, “Do professional traders exhibit loss realization aversion?” working paper, Texas Christian University.


