THREE ESSAYS IN APPLIED ECONOMICS

by

AUSTIN C. SMITH

B.S., Bentley University, 2010

M.A., University of Colorado, 2012

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Economics

2015

This dissertation entitled:

Three Essays in Applied Economics

By **Austin C. Smith**

has been approved for the Department of Economics

Professor **Brian Cadena**, Co-chair

Professor **Jonathan Hughes**, Co-chair

Date: April 9, 2015

The final copy of this thesis has been examined by the signatories, and we find that the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

ABSTRACT

Smith, Austin C. (Ph.D., Economics)
Three Essays in Applied Economics
Thesis directed by Assistant Professors Brian Cadena and Jonathan Hughes

This dissertation examines various topics in applied economics, with particular attention to policy evaluation and to how workers respond to incentives. In the first chapter, I examine the impact of Daylight Saving Time (DST) on fatal automobile accidents. Despite mounting evidence that DST fails in its primary goal of saving energy, some form of DST is still practiced by over 1.5 billion people in over 60 countries. I demonstrate that DST imposes high social costs on Americans, specifically, an increase in fatal automobile crashes. DST alters fatal crash risk in two ways: disrupting sleep schedules and reallocating ambient light from the morning to the evening. I find that shifting ambient light reallocates fatalities within a day, while sleep deprivation caused by the spring transition increases risk. The increased risk persists for the first six days of DST, causing a total of 302 deaths at a social cost of $2.75 billion over the 10-year sample period.

In the second chapter, my coauthor and I examine the impact of alternative compensation structures on worker performance. We use a unique panel dataset that tracks workers as they are able to adjust how they are paid for performance between a pure variable payment scheme based on performance in a rank-order tournament and a hybrid structure with a fixed fee and tournament component. We find that performance is best under the pure variable payment structure, with performance under the hybrid structure improving as more weight is given to the variable component of pay. Importantly, the incentives created by the variable payment scheme have an economically meaningful impact in a setting where output is highly variant.

In the final chapter, I investigate how professional online poker players respond to shocks in expected wages. Using variation in expected wages created by the amateur portion of the player pool, I find that the top pros respond on the extensive margin, playing more often when expected wages are high, and on the intensive margin, playing more tables simultaneously. In contrast, weaker pros respond only on the intensive margin.

ACKNOWLEDGEMENTS

Thank you to my advisors, Brian and Jon, for teaching me how to become a producer of economics. Thank you to Greg, for being the best co-author one could hope for. Thank you to my parents and siblings, who encouraged me to pursue graduate school. Most importantly, thank you to my wife. Kendra, thank you for putting up with me during the stressful times, celebrating with me during the exciting times, and encouraging me at the low points of the research roller coaster. I could not have done this without you.

Contents

I Spring Forward at Your Own Risk: Daylight Saving Time and Fatal Vehicle Crashes **1**

1 Introduction **1**

2 Daylight Saving Time in the US **6**

3 Mechanisms **7**
3.1 Sleep Mechanism 8
3.2 Light Mechanism 9

4 Data **10**
4.1 FARS 10
4.2 Other Data Sources 11

5 Empirical Strategy **13**
5.1 Regression Discontinuity (RD) Methods 13
5.2 Day-of-Year Fixed Effects 17

6 Results **19**
6.1 Spring RD Design 19
6.2 Fall RD Design 21
6.3 Mechanisms 23
6.3.1 Light 24
6.3.2 Sleep 25
6.4 Fixed Effects Model 28
6.5 Alternative Explanations 33
6.6 Geographical Heterogeneity 34

7 Conclusion **37**

II Alternative Compensation and Performance: Evidence From Staking in Online Poker **39**

8 Introduction **39**

9 Online Poker Tournaments and the Market for Staking **43**
9.1 Online Poker Tournaments 43
9.2 The Market for Staking 45

10 Data **47**

11 Empirical Strategy **50**
11.1 Monetary Outcomes 52
11.2 Finish Position 52

12 Results **54**
12.1 Main Results 54
12.1.1 Monetary Outcomes 54
12.1.2 Finishing Rank Outcomes 57
12.2 Addressing Issues of Selection Bias 60
12.2.1 Staking Timeline 61
12.2.2 Percent Staked 62
12.2.3 Staking Gap 67

13 Conclusion **70**

III Scanning the Tables: Do Professional Online Poker Players Respond to Earning Conditions? **72**

14 Introduction **72**

15 Online Poker Cash Games **74**

16 Data Description **75**
16.1 Defining Professionals 79

17 Empirical Framework **81**
17.1 Predicting "Skill" of Amateurs 83
17.2 Labor Supply Responses to Earning Conditions 85

18 Results **86**
18.1 Main Results 86
18.2 The Start and Quit Decisions 89
18.3 Robustness - Addressing the Exogeneity Assumption 93
18.4 Aggregate Daily Measures 94

19 Conclusion **100**

List of Tables

1 RD estimates of the impact of entering DST on fatal crashes 20
2 RD estimates of the impact of entering DST on fatal crashes - additional robustness 22
3 RD estimates of the impact of entering DST on fatal crashes - global polynomial regressions 22
4 RD estimates of the impact of leaving DST on fatal crashes 23
5 RD estimates of the influence of ambient light on fatal crashes when leaving DST 25
6 RD estimates of the influence of sleep loss on fatal crashes 28
7 FE estimates of the impact of DST on fatal crashes 31
8 RD estimates of the impact of entering DST on fatal crashes, by county risk level 35
9 RD estimates of the impact of entering DST on fatal crashes - geographical impacts 36
10 Summary Statistics 49
11 Monetary Outcomes 56
12 Finishing Position Outcomes 59
13 Monetary Outcomes - Percent Staked (Restricted Sample) 64
14 Finishing Rank Outcomes - Percent Staked (Restricted Sample) 66
15 Monetary Outcomes (Staking Gap and Percent Requested; Restricted Sample) 69
16 Summary Statistics 80
17 The impact of earning conditions on whether and how much to play . . . 87
18 The impact of earning conditions on when to start and quit a poker session 91
19 Robustness - Controlling for Amateur Volume 95
20 Aggregated Daily Observations 97

List of Figures

1 The Influence of Daylight Saving Time on Ambient Light 2
2 Frequency of Fatal Crashes by Hour 4
3 Fatal Crashes Around the Spring Transition 12
4 Spring Residual Plot 15
5 Fall Residual Plot 16
6 Variation in DST Coverage - Spring 18
7 Reallocation of Fatal Crashes (Fall Transition) 26
8 Spring Residual Plot – Six Day Sleep Impact 27
9 Spring Residual Plot – Least Light Impacted Hours 29
10 The Impact of Staking on Poker Tournament Prizes 44
11 Typical Staking Transaction 46
12 Distribution of *finishpercentile*: Staked versus Unstaked 54
13 Residual Plot of *finishpercentile* Around the First Staking Incident 62
14 Dataset Characteristics and Matching 78
15 Dataset Alignment 79
16 Distribution of Weekly Hours Played 81
17 Histogram of Residual Variation in *AmExpCont* 99

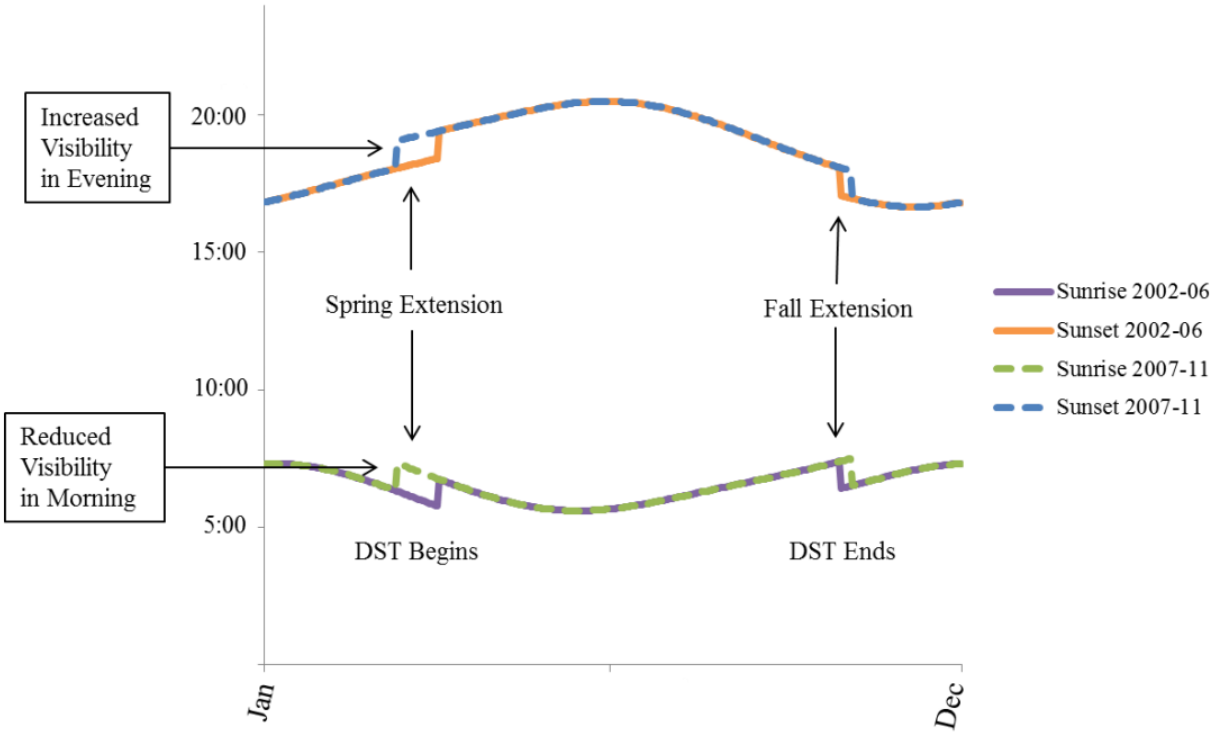
Chapter I
Spring Forward at Your Own Risk: Daylight Saving Time and Fatal Vehicle Crashes

1 Introduction

Daylight Saving Time (DST) in the US was originally implemented as a wartime measure to save energy and was extended as part of the Energy Policy Act of 2005. However, recent research demonstrates that DST does not save energy and could possibly increase energy use (Kellogg and Wolff, 2008; Kotchen and Grant, 2011). Despite mounting evidence that DST fails in its primary goal, some form of Daylight Saving Time is still practiced by over 1.5 billion people globally. In this paper I demonstrate that DST imposes high social costs on Americans, specifically, an increase in fatal automobile crashes. Employing three tests to differentiate between an ambient light or sleep mechanism, I show that this result is most likely due to sleep deprivation caused by the spring transition and the result implies additional costs of DST in terms of lost productivity nationwide.

The procedure for DST is well characterized by the phrase "spring-forward, fall-back." Each year on the spring transition date, clocks are moved forward by one hour, from 2 a.m. to 3 a.m. The process is then reversed for the fall transition with clocks "falling back" from 2 a.m. to 1 a.m. This alters the relationship between clock time and solar time by an hour, effectively moving sunlight from the morning to the evening (see Figure 1). The procedure was first suggested by George Vernon Hudson, an entomologist who wanted more light in the evenings to pursue his passion of collecting insects (Hudson, 1895). While the policy was first used during World Wars I and II, it has since become

Figure 1: The Influence of Daylight Saving Time on Ambient Light



Note: The sunset and sunrise times are for St. Louis Missouri, the nearest major city to the population center of the US.

a peacetime measure. In all instances, the rationale has been that aligning sunlight more closely with wakeful hours would save energy used for lighting.[1] However, as Hudson's personal motivation for the policy suggests, DST has many impacts on practicing populations.

This paper focuses on a major side-effect of DST, its impact on fatal vehicle crashes. DST alters the risk of a fatal crash in two ways: disrupting sleep schedules and reallocating ambient light from the morning to the evening. With an average of over1 39,000 annual fatalities, motor vehicle crashes are the number one cause of accidental death in the US (CDC, 2005-2010). Given the large base level of fatalities, even a small change in

[1]DST is often mistakenly believed to be an agricultural policy. In reality, farmers are generally against the practice of DST because it requires them to work for an extra hour in the morning, partially in darkness, to coordinate with the timing of markets (Prerau, 2005).

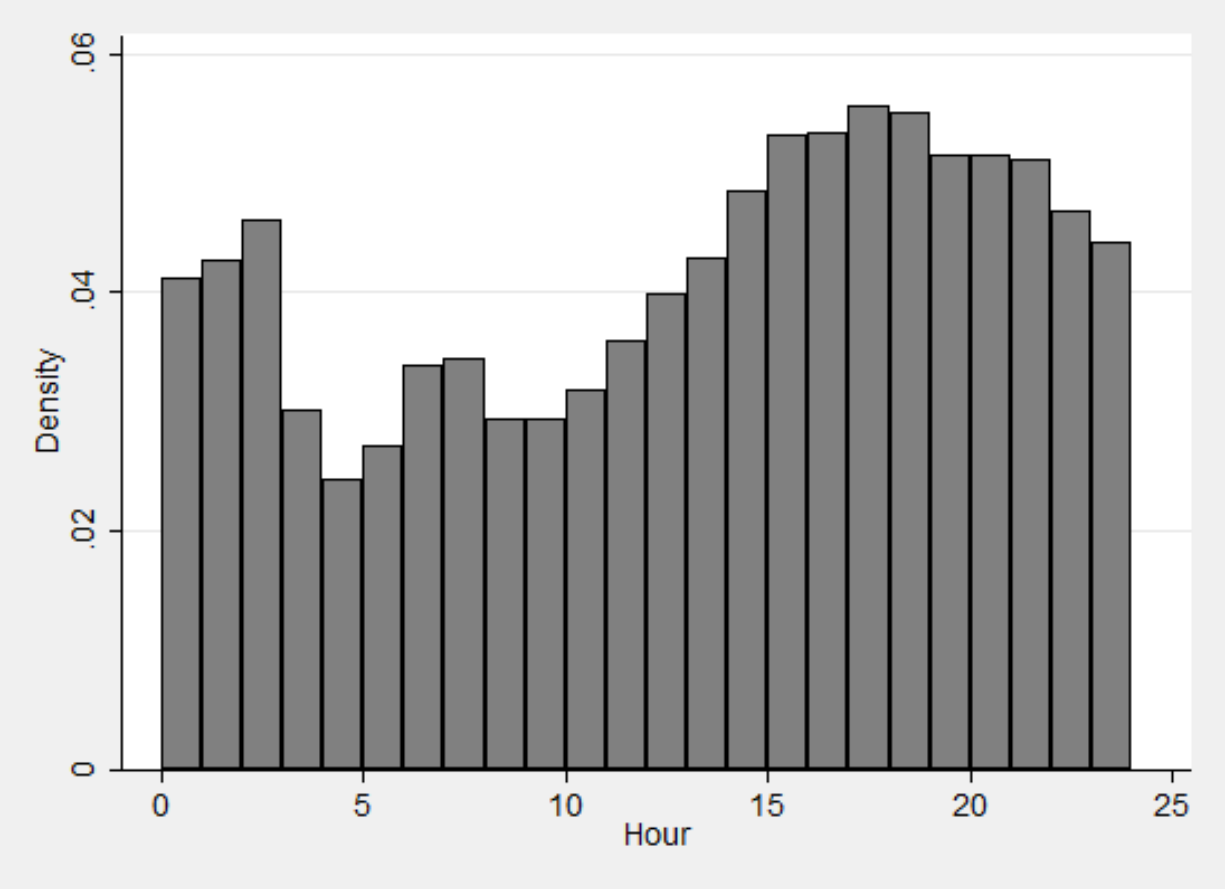
fatal crash risk is a potentially large killer. I identify the impact of DST on fatal crashes by taking advantage of (i) detailed records of every fatal crash occurring in the United States from 2002-2011; (ii) the discrete nature of the switch between Standard Time and Daylight Saving Time; and (iii) variation in the dates covered by Daylight Saving Time, created primarily by a 2007 policy change. I employ two different identification strategies. First, I use a regression discontinuity (RD) design that examines changes in daily crash counts immediately before and after DST transitions. Second, to measure the duration of impact, I use a day-of-year fixed effects (FE) model that is identified by dates that are covered by DST in some years but Standard Time in other years. In both specifications I find a 5.4-7.6% increase in fatal crashes immediately following the spring transition. Conversely, I find no impact following the fall transition when no shock to sleep quantity occurs.[2] To address the possibility that some other unobserved factor related to the transition dates is driving this result, I impose the pre-2007 transition dates on data from 2007-2011 and the current transition dates on data from 2002-2006 and find no impact of these dates when not associated with a policy change. I then examine the relative contribution of each DST mechanism.

Daylight Saving Time impacts practicing populations through two central channels. First, it creates a short-term disruption in sleeping patterns following the spring transition. Harrison (2013) surveys the sleep literature and finds that "increased sleep fragmentation and sleep latency" caused by the 23-hour spring transition date "present a cumulative effect of sleep loss, at least across the following week." Second, DST alters the relationship between clock time and solar time by an hour, creating darker mornings and lighter evenings than would be observed under Standard Time (see Figure 1).[3] Even

[2]Barnes and Wagner (2009) find that Americans sleep 40 minutes less on the night of the spring transition, but experience no significant change in sleep quantity on the fall transition.

[3]Since fatal crashes are more prevalent in the evening (Figure 2), it is possible that transferring light from a lower risk morning period to a higher risk evening period could lead to a net reduction in fatal crashes.

Figure 2: Frequency of Fatal Crashes by Hour



Note: Histogram uses all fatal crashes from 2002-2011 in the contiguous US except Arizona and Indiana.

this one hour shift in light can have major consequences; Doleac and Sanders (2013) find that increased ambient light in evenings reduces crime while Wolff and Makino (2013) suggest that it increases time devoted to exercise.

To parse out these mechanisms and determine what portion of the increase in fatal crashes is due to sleep loss versus reallocating light, I run three primary tests. These tests exploit differential timing in when each mechanism is active, both within and across days. First, I isolate the light mechanism by examining only the fall transition.[4]

[4]Americans do not sleep a significant amount more on the fall transition date despite receiving an extra hour in the middle of the night (Barnes and Wagner, 2009).

Then, I look at the difference between aggregate estimates in the fall (only the light mechanism) and spring (light and sleep mechanism) to determine the net impact of the sleep mechanism. Second, I isolate the sleep mechanism in the spring by examining a subsample of hours furthest from sunrise and sunset. These hours are least impacted by the light mechanism and a drowsy driver is presumably more at risk throughout the entire day, even in hours of full light or full darkness. Third, I compare the sleep impacted days of DST (up to the first two weeks) to the remainder of DST with common support.[5] All three tests suggest that the sleep deprivation is driving the increase in fatal crashes.

My preferred specification reveals a 6.3% increase in fatal crashes, persisting for six days following the spring transition. Over the 10-year sample period, this suggests the spring transition is responsible for a total of 302 deaths at a social cost of $1.2 to $3 billion, underscoring the huge costs of even minor disruptions to sleep schedules given the current sleep-deprived culture in the US.[6,7] The total costs of DST due to sleep deprivation could be orders of magnitude larger when worker productivity is considered (Wagner et al., 2012; Kamstra, Kramer, and Levi, 2000).[8]

This finding is timely, given the recent empirical research suggesting that DST does not reduce energy demand. Kellogg and Wolff (2008) use a natural experiment in Australia where DST was extended in some states to accommodate the Sydney Olympics. They find that while DST reduces energy demand in the evening, it increases demand in the morning with no significant net effect. Kotchen and Grant (2011) make use of quasi-experiment in Indiana where some Southern Indiana counties did not practice DST until

[5] Common support refers to dates that are DST in some years and Standard Time in others.

[6] Social cost is based on Kniesner et al. (2012) value of a statistical life range of $4 to $10 million.

[7] Nearly 30% of American adults reported sleeping less than 6 hours per day in 2005-2007 according to a National Center for Health Statistics survey.

[8] There has been surprisingly little empirical research on the effects of sleep on worker productivity. Although fatal crashes are an extreme measure of productivity, driving is a behavior engaged in by over 90% of American workers (Winston, 2013) and the increase in fatal crashes suggests that sleep loss likely reduces productivity.

2006. Their work suggests that DST could actually increase residential energy use, as increased heating and cooling use more than offset the savings from reduced lighting use. For a failed energy policy to be justified from a welfare standpoint, the social benefits must outweigh the social costs. In this paper, I find a significant mortality cost that must be weighed against any perceived benefits of DST.

The remainder of the paper is organized as follows. The next section provides a brief background of DST in the US. Section 3 details the mechanisms through which DST influences crash risk, including reviewing existing evidence of the impact of DST on vehicle crashes. Section 4 introduces the data, highlighting the visual discontinuity in raw crash counts at the spring transition. Section 5 describes the RD and FE identification strategies, outlining the requirements for causal estimates. Section 6 presents results, including those that differentiate between the sleep and light mechanisms, and explores alternative explanations. Section 7 concludes with a brief summary and further remarks about the implications for DST as a policy.

2 Daylight Saving Time in the US

Daylight Saving Time has been a consistent feature in most US states since the Uniform Time Act of 1966.[9] This legislation allowed states to determine whether they practiced DST, but set uniform start and stop dates for any practicing states. Since 1966, Congress has twice made lasting changes to the DST transition dates, most recently as part of the Energy Policy Act of 2005. Starting in 2007, DST begins on the second Sunday of March and continues until the first Sunday of November, a 3-4 week extension in the spring and a 1 week extension in the fall.

[9] Among the contiguous US, all states but Arizona and parts of Indiana have practiced DST since 1973.

Figure 1 illustrates the impact of DST on sunrise and sunset times throughout the year and highlights the 2007 extension. On the spring transition date, clocks skip forward from 2 to 3 a.m. pushing sunrise and sunset times back by one hour. In the fall, the process is reversed as clocks are adjusted back by an hour to facilitate the return to Standard Time. The 2007 extension to DST altered these transition dates and created an additional range of dates that are DST in some years and Standard Time in others.[10] In the next section, I discuss the primary mechanisms through which DST could influence fatal crash risk and how I disentangle the relative contributions of each.

3 Mechanisms

There are two mechanisms through which Daylight Saving Time could impact fatal crash risk. First, there is sleep loss associated with the spring transition when one hour in the middle of the night is skipped. Since sleep is a key factor in alertness and control (Smith, McEvoy, and Gevins, 2002), this sleep deprivation likely reduces driving safety. In a study of 400 U.S. Army soldiers, Legree et al. (2003) find a correlation of 0.20 between driver at fault accidents and self reported insufficient sleep. Second, DST shifts the mapping of solar time to clock time by an hour, reallocating sunlight between the morning and the evening. Ambient light reduces fatal crash risk (Fridstrom et al., 1995; Sullivan and Flannagan, 2002), and this reallocation of light within a day creates riskier morning driving conditions and less risky evening driving conditions during DST.[11] I next discuss each mechanism individually, outlining its likely effect on fatal crashes and reviewing existing evidence of its impact through DST.

[10] Since transition rules are based on moving dates (e.g. the second Sunday of March ranges from 3/8 to 3/14) there is variation in start and end dates even within a particular transition rule.

[11] When switching out of DST in the fall, the mornings become less risky and evenings more risky than under DST.

3.1 Sleep Mechanism

The spring transition into DST is facilitated by clocks jumping forward from 2 a.m. to 3 a.m. on the transition date. This creates a 23-hour transition day, rather than the standard 24-hour days people are accustomed to. While this "missing" hour could be cut from work or leisure time, Barnes and Wagner (2009) find that Americans make up the majority of the missing time by sleeping less. Using the American Time Use Survey, they find Americans sleep an average of 40 minutes less on the night of the spring transition. Depending on the individual, this transition could impact sleep patterns for anywhere from two days to two weeks (Valdez et al., 1997) with an average of about one week (Harrison, 2013).

In the fall, the opposite scenario occurs with a 25-hour transition day. However, in this case, Americans use very little of the extra hour for sleep, sleeping a statistically insignificant extra 12 minutes (Barnes and Wagner, 2009). This creates variation in treatment status for the sleep mechanism. The spring transition is treated (sleep loss), while the fall transition is untreated (insignificant change to sleep quantity).[12]

Previous research on the sleep impact of DST on vehicle crashes has been mixed. Coren (1996) and Varughese and Allen (2001) find an increase in crashes on the Monday following the spring transition into DST, while Sood and Ghosh (2007) and Lahti et al. (2010) suggest no effect. By focusing on one day, these tests can lack power and often cannot rule out a wide range impacts. In contrast to these studies, I gain statistical power by testing for a longer term sleep impact consistent with recent literature on sleep disruptions.

Additionally, these previous studies use data centered in 1992, 1985, 1987 and 1994 respectively. Average sleep quantity has been on the decline in the US, a phenomenon

[12]Sexton and Beatty (2014) also find significant sleep loss associated with the spring transition but no significant change in the fall.

also seen in the lower tail of the distribution. According to the National Sleep Foundation, the percentage of Americans averaging less than 6 hours of sleep has risen from 12% in 1998, to 20% in 2009. My data spans 2002-2011 and should generate a more up to date measure of the impact of sleep loss given the current sleep patterns in the US.

3.2 Light Mechanism

Despite strong evidence suggesting the importance of ambient light in fatal crash risk, the implication for net crashes due to Daylight Saving Time remains unclear. DST does not alter the amount of light in a day, it simply reallocates it between the morning and the evening. Since fatal crashes are more prevalent in the evening (Figure 1), it is possible that transferring light from a lower risk morning period to a higher risk evening period could lead to a net reduction in fatal crashes.

Previous studies by Ferguson et al. (1995) and Broughton, Hazelton, and Stone (1999) examine the light mechanism by estimating the impact of ambient light on fatal crash risk directly, and then simulating the impact of imposing DST light levels on the rest of the year. Both studies suggest a reduction in fatal crashes through this mechanism.[13] However, the simulation in Ferguson et al. (1995) uses a single measure of the impact of light on crash risk. This generates a biased estimate of the life saving potential of DST if ambient light interacts with other risk factors such as driver alertness, or type of trip (work versus leisure) both of which are likely to vary from morning to evening driving. Further, simulation requires assumptions about driver behavior under counterfactual hours of light.

As an alternative to these simulation methods, I use empirical techniques to estimate

[13] Sood and Ghosh (2007) also find a reduction in crashes which they attribute to the light mechanism. However, they analyze only the spring transition and results are sensitive to the time frame analyzed and choice of control group.

the effect directly. First, I focus on the fall transition as a clean estimate of the light mechanism because it is not afflicted by any significant shock to sleep. Then, I examine the spring following the first two weeks of DST, when the sleep mechanism should no longer be active.

4 Data

4.1 FARS

For vehicle fatality data, I use the Fatality Analysis Reporting System (FARS), compiled by the National Highway Traffic and Safety Administration. These data contain a record of every fatal crash occurring in the United States since 1975, including exact time and location of the accident. I focus on recent crashes, from 2002-2011, allowing for five years on either side of the 2007 DST extension. Consistent with other DST literature, my sample is the continental US excluding Arizona and Indiana because at least part of those states did not practice DST consistently over the entire sample time frame.[14] Since the initial Sunday of DST is 23 hours long, whereas other days are 24 hours long, I adjust the crash count by counting the 3-4 a.m. hour twice, using it as a proxy for the missing 2-3 a.m. hour. For the 25-hour fall transition date, I divide the fatalities occurring from 2-3 a.m. by two, because this hour occurred twice.[15]

My dependent variable in all specifications is the natural log of the number of fatal crashes occurring on a given day at the national level. I aggregate to the national level due to the relative rarity of fatal crashes. There are roughly 100 fatal crashes

[14] Less than 1% of the remaining observations are dropped due to missing or inaccurate time of day.

[15] I also use two alternative corrections, multiplying crashes on the spring transition date by 24/23rds and those on the fall transition date by 24/25ths, or simply dropping the transition dates from the sample. Results are robust to both methods.

per day across the entire US and the mode for daily crashes at the state level is zero. Aggregating allows me to gain statistical power and smooths out potential confounders such as weather which could drive results in some states or even regions, but likely not the entire US.

Figure 3 plots the total number of fatal crashes occurring in the weeks surrounding the spring transition into DST. There is a clear break in the seasonal trend of fatal crashes, occurring right at the spring transition.[16] This provides suggestive evidence that the spring transition is associated with a short term increase in fatal crashes. My initial estimation strategy (RD) formally tests for this discontinuity.

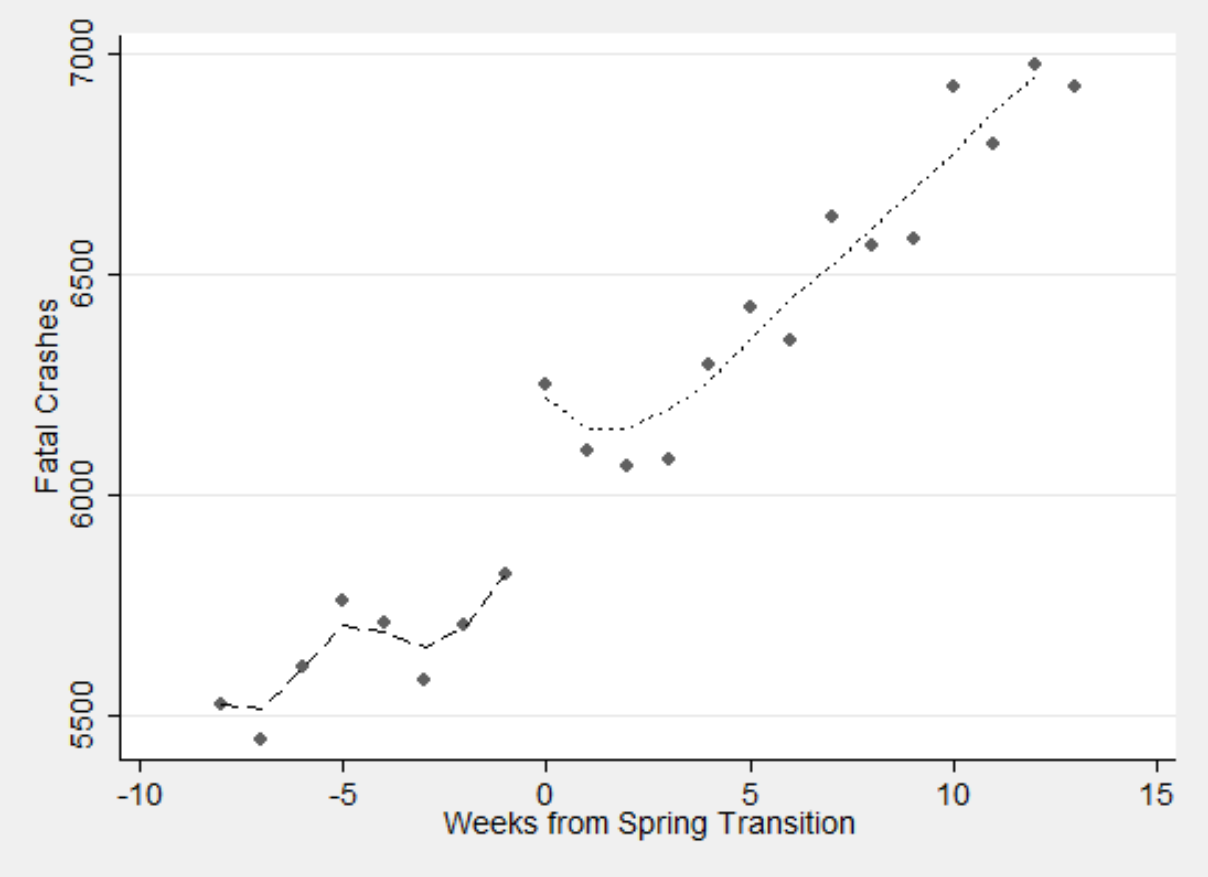
If complete data were available for less severe crashes, it could be analyzed in the same identification framework I propose. However, many states do not maintain a uniform database of these less severe crashes and the potential for reporting bias and less rigorous redundancy checks for non-fatal crashes make these data less reliable. Considering only fatal crashes is likely a lower bound on the impact of DST on all automobile crashes.

4.2 Other Data Sources

Fridstrom et al. (1995) find "exposure to risk" or Vehicle Miles Traveled (VMT) to be the most important predictor of fatal crash counts. Unfortunately, daily VMT data does not exist at the national level. As such, I use VMT data from Caltrans' Performance Measurement System (PeMS) to examine whether adjustments to VMT are driving my results. To the extent that VMT on this subset of roads is representative of US driving patterns, this provides a useful test. In the national sample, I use weekly gasoline prices from the U.S. Energy Information Administration and the value of the S&P 500 index

[16] The seasonal trend is largely due to a similar seasonal increase in vehicle miles traveled.

Figure 3: Fatal Crashes Around the Spring Transition



Notes: Each point represents the total number of fatal crashes occurring during that week from 2002-2011. Smoothed lines are results of locally weighted regression.

to help control for fuel prices and driving patterns.

5 Empirical Strategy

5.1 Regression Discontinuity (RD) Methods

The goal of the empirical analysis is to identify the impact of DST on fatal motor vehicle crashes. To perform this analysis, I use a regression discontinuity design that exploits the discrete change from Standard Time to DST. Every year on the spring cutoff date, clock time is altered by one hour. If there is a significant impact of DST on fatal crashes, there should be a shock to the number of fatal crashes from just before to just after the transition. Measuring the discontinuity occurring at the policy transition provides an estimate of the policies immediate impact.

My preferred specification uses local linear regression, as it has been shown to perform better in RD settings than high order polynomials of the running variable (Gelman and Imbens, 2014).[17] To eliminate persistent day-of-week effects (e.g. crashes are higher on weekends than weekdays) and long-term time trends, I first demean the logged crash counts by day-of-week and year. Then, I use the standard RD specification with the demeaned crash data. The estimation equation is seen below:

$$\ln Fatals_{dy} = \beta_0 + \beta_1 DST_{dy} + \beta_2 DaysToTran_{dy} + \beta_3 DST_{dy} * DaysToTran_{dy} + \varepsilon_{dy} \quad (1)$$

DST_{dy} is an indicator equal to one if day d in year y falls under Daylight Saving Time and $DaysToTran_{dy}$ is the running variable, measuring time in days before and after the DST transition. $DaysToTran_{dy}$ is centered at the transition date in each year, the

[17] Results using a global polynomial are qualitatively identical and are available in Table 3.

first Sunday of April in 2002-2006 and the second Sunday of March in 2007-2011. The coefficient of interest, β_1, is the aggregate effect of DST on vehicle fatalities at the cutoff date.[18]

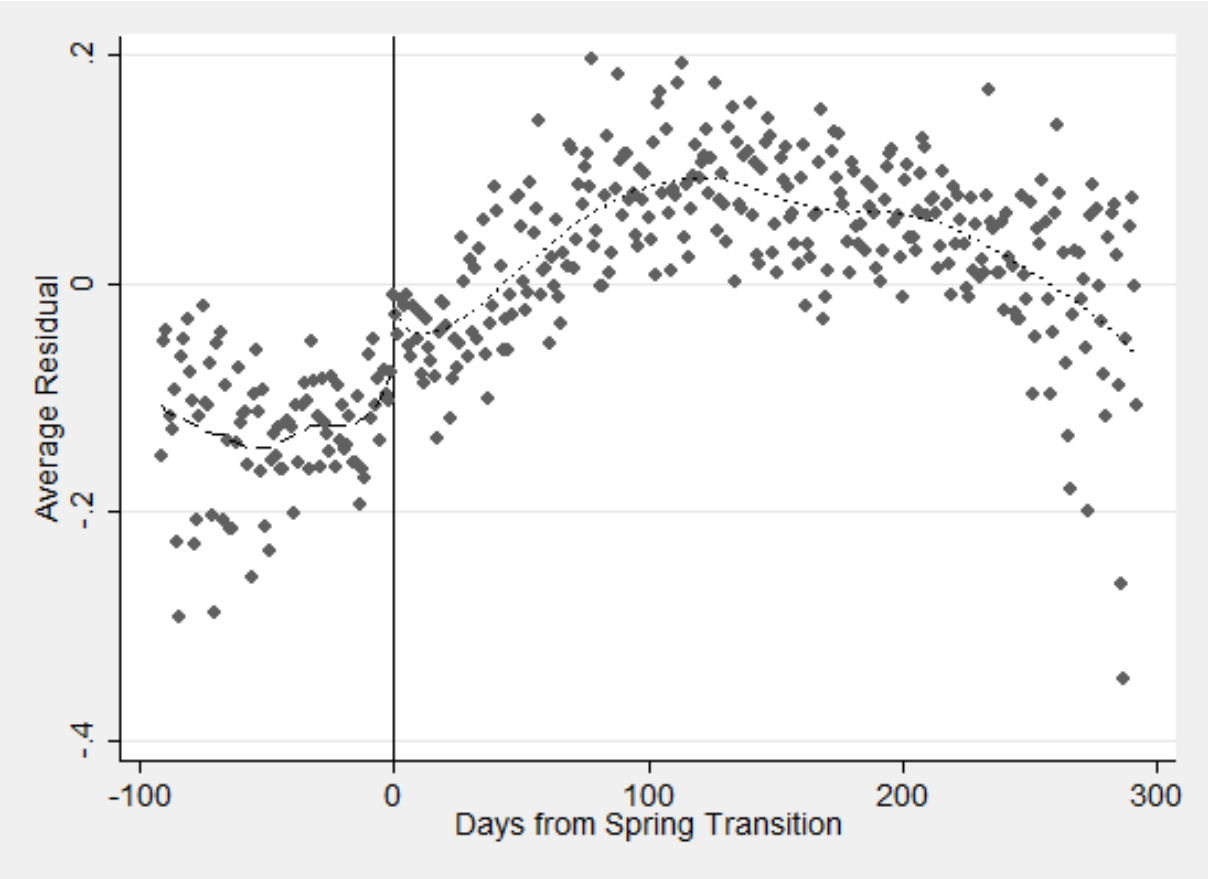
My baseline specification uses Calonico, Cattaneo, and Titiunik's (2012) optimal bandwidth selector to determine how many days to use on either side of the DST transition and a uniform kernel. As Imbens and Lemieux (2008) argue, there is little practical benefit to other weighting schemes as they are primarily indicative of sensitivity to the bandwidth choice. For robustness I include results using alternative bandwidth selectors and Epanechnikov and triangular kernels.

In this context, a consistent estimate requires that conditional on day of the week and year, the treated and untreated number of fatal car crashes must vary continuously with date around the transition. Stated differently, if all other factors affecting fatal crash risk, besides DST, are continuous at the transition date, the RD design will provide consistent estimates of the effect of DST. Figures 4 and 5 begin to speak to this assumption, providing visual evidence that after demeaning the data, fatal crashes vary smoothly across a year. In Section 6.5, I directly test for discontinuities in other factors that impact crash risk.

The Energy Policy Act of 2005 allows me to further probe the robustness of my RD estimates in a difference in discontinuities placebo test. The new March transition date went into effect in 2007 and should have no impact in previous years. Likewise, the old April transition date should not impact crashes in 2007-2011. By looking for a discontinuity using these placebo transition dates, I can test whether these dates are typically associated with a change in fatal crashes, unrelated to DST. I apply the analogous procedure to the fall transition.

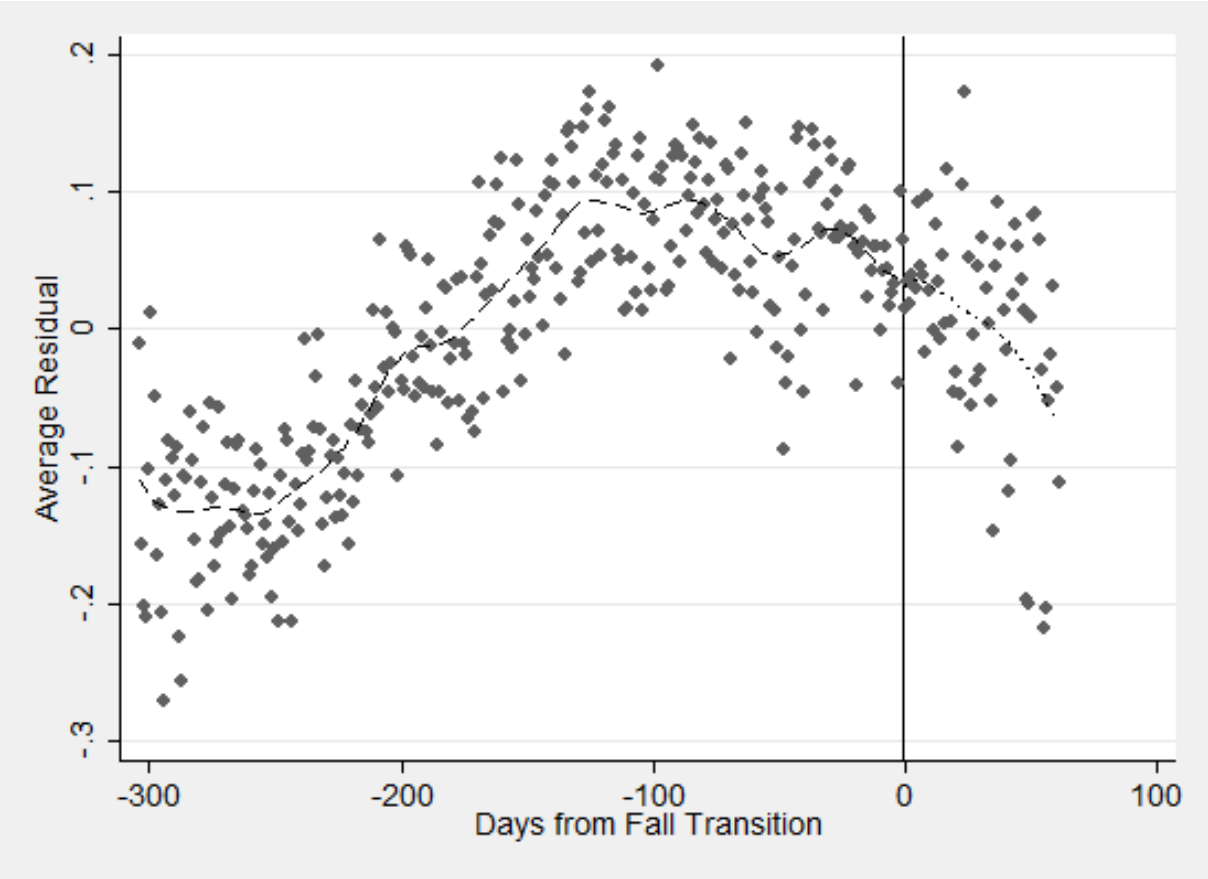
[18] I refer to this as the aggregate impact, because it does not yet disentangle the DST mechanisms.

Figure 4: Spring Residual Plot



Notes: The residuals are generated from a regression of ln(fatal crash count) on day-of-week and year dummies. Each point is the average of all residuals for that date relative to the spring transition. Fitted lines are results of locally weighted regression. Greater variability on the ends is largely due to these average residuals being formed by only 5 observations rather than 10 towards the middle. This is a product of the 2007 DST extension; in 2002-2006 there are about 14 weeks before the spring transition but in 2007-2011 about 11.

Figure 5: Fall Residual Plot



Notes: The residuals are generated from a regression of ln(fatal crash count) on day-of-week and year dummies. Each point is the average of all residuals for that date relative to the fall transition. Fitted lines are results of locally weighted regression. Greater variability on the ends is largely due to these average residuals being formed by only 5 observations rather than 10 towards the middle. This is a product of the 2007 DST extension; in 2002-2006 there are about 9 weeks following the fall transition but in 2007-2011 about 8.

5.2 Day-of-Year Fixed Effects

While the RD design provides a measure of the causal impact of DST on fatal crashes at the transition date, it is more limited in estimating longer term impacts. To empirically estimate these longer lasting effects, I leverage variation in the coverage of Daylight Saving Time created by both the 2007 extension and the DST cutoff rules. From 2002-2006 the time period between the second Sunday of March and the first Sunday of April was part of Standard Time. The Energy Policy Act of 2005 extended DST to cover this 3-4 week period in 2007-2011. This creates a range of dates that are DST in some years and Standard Time in other years. The cutoff rule further expands the number of "switching days". Consider the current decision rule where DST begins on the second Sunday in March. The start date has varied from the 8th to the 14th of March depending on the year.[19] Figure 6 shows days of the year that fall under both DST and Standard Time during the spring and their frequency under each regime. During the fall there is a similar, but smaller, region of switching dates because the fall transition date was only pushed back by one week.

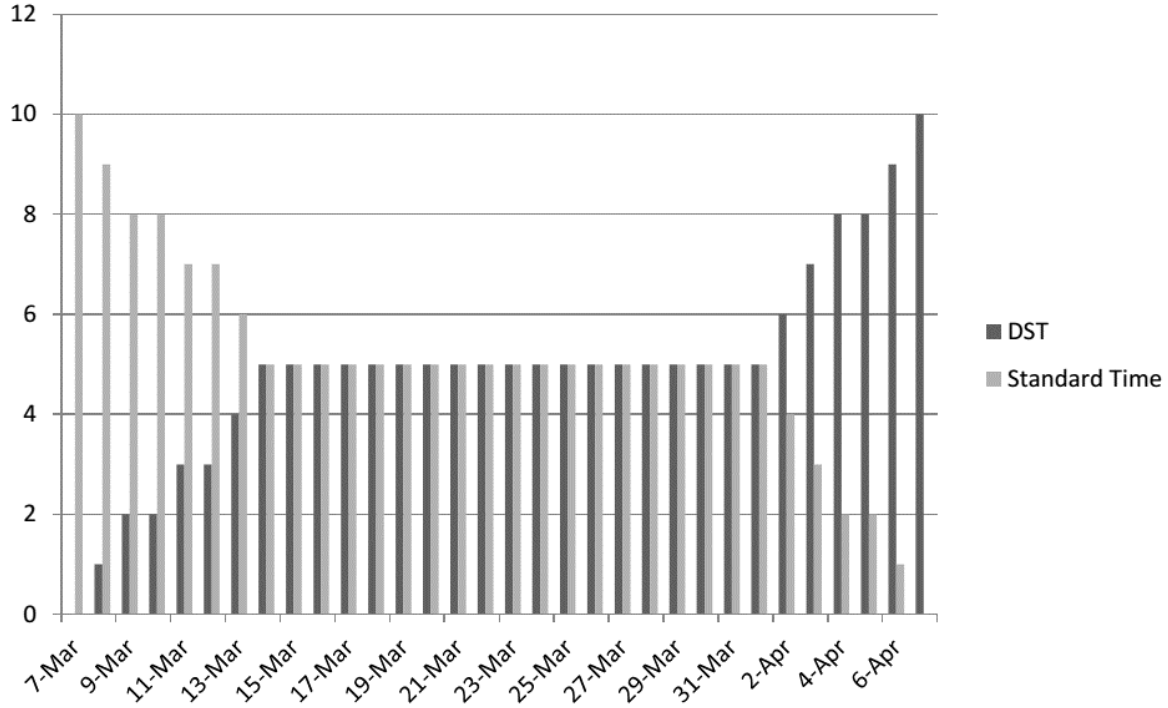
Moving to a fixed effects framework, I run the following specification to take advantage of this variation in DST assignment:

$$\begin{aligned} \ln Fatals_{dy} &= \beta_0 + \beta_1 SpDST_{dy} + \beta_2 FaDST_{dy} + DayofYear_d \\ &\quad + DayofWeek_{dy} + Year_y + V_{dy} + \varepsilon_{dy} \end{aligned} \tag{2}$$

$DayofYear_d$ is a separate dummy for each day of the year, flexibly controlling for the

[19] For example, March 11th is Standard Time in 2002-2006, 2010 and 2011 but is DST in the years 2007-09.

Figure 6: Variation in DST Coverage - Spring



impact of seasonality on fatal crashes.[20] $DayofWeek_{dy}$ and $Year_y$ are day-of-week and year dummies respectively. V_{dy} is a vector of controls used in some specifications, including gasoline prices, the value of the S&P 500 index and non-stationary holidays. $SpDST_{dy}$ is an indicator equal to one if the date falls under DST and is covered by the range of spring switching dates (March 8th - April 7th). Analogously, $FaDST_{dy}$ is an indicator equal to one if the date falls under DST and is covered by the range of switching dates in the fall (Oct 25th - Nov 7th). These are the coefficients of interest and are interpreted as the average effect of DST on fatal crashes over the "switching" dates in that season.

Note, that β_1 here is a different parameter from what is found using the RD design. Regression discontinuity estimates the effect of DST right at the spring transition,

[20] I create dummies for each month/day combination (e.g. an August 25th dummy). This is slightly different than creating a dummy for the 100th day of the year, because leap day would cause August 25th for most years to be matched with August 24th for 2004 and 2008. I use the month/day method as it better aligns with holidays and generates more conservative estimates.

whereas the fixed effects specification measures the average effect of DST over all dates that are sometimes DST and sometimes Standard Time during the spring. If DST only creates a short-run effect through sleep deprivation, this should be picked up in the RD, but would be averaged out across the full range of switching dates when using the fixed effects model. Likewise, β_2 is the average effect of DST across the roughly two weeks of fall switching dates, rather than the effect of leaving DST in the fall.

Beyond identifying the average effect of DST across the range of switching dates, this specification can aid in disentangling the mechanisms. I isolate the light mechanism in the spring, by focusing only on dates at least two weeks following the transition, at which time any sleep impact should have dissipated. Comparing this light impact to the initial impact from light and sleep provides another measure for just the sleep impact.

6 Results

6.1 Spring RD Design

Figure 4 illustrates the regression discontinuity strategy for estimating the impact of DST on fatal crashes. The average residuals from a regression of log(daily fatal crash count) on day-of-week and year dummies are plotted, centered by the spring transition date. The plot follows a gradual arc demonstrating the seasonal pattern in fatal crashes, where crashes rise from winter lows, peaking in late summer before dropping again through the fall. If DST has an impact on fatal crashes, this should be evident in a trend break right at the transition date. Visually, there is a short-term spike in fatal crashes before the residuals resume the seasonal trajectory.

Table 1: RD estimates of the impact of entering DST on fatal crashes

	(1)	(2)	(3)	2002-2006 (4)	2007-2011 (5)	Placebo (6)
DST	0.0631**	0.0536**	0.0756***	0.0682**	0.0949	-0.0174
	(.0309)	(.0215)	(.0218)	(.0341)	(.0583)	(.0278)
Bandwidth	CCT	IK	CV	CCT	CCT	CCT
# days left	18	41	57	20	12	20
# days right	19	42	58	21	13	21

Dependent Var: Log fatal crashes; all specs use day-of-week and year dummys, a first order polynomial and a uniform kernel. DST is the estimate of the discontinuity in fatal crashes that occurs immediately following the spring transition into DST. Placebo assigns the current March transition date to 2002-2006 data and the old April transition date to the 2007-2011 data. CCT refers to the bandwidth selector of Calonico, Cattaneo, and Titiunik (2012); IK is Imbens and Kalyanaraman (2012); CV is the cross-validation method of Ludwig and Miller (2007). Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 1 shows the corresponding regression estimates.[21] The spring transition into DST is associated with a 6.3% increase in fatal crashes. This result persists using the bandwidth selectors of Imbens and Kalyanaraman (2012) and the cross-validation method of Ludwig and Miller (2007) seen in columns 2 and 3 respectively. To test whether the increase is due to one particular transition rule, I split the data into an early subsample (2002-06) that was subject to the April transition, and a late subsample (2007-2011) that is subject to the current March transition. While cutting the sample in half reduces precision, both time periods experience similar increases in fatal crashes at the transition.[22]

To address the possibility that both transition dates are associated with an increase in fatal crashes, unrelated to DST, I run the following placebo test in column 6. I assign the current transition date to 2002-2006 data and the old transition date to the

[21] Clustering by week or year tends to *decrease* standard errors as the shocks are negatively correlated, so I report the more conservative uncorrected standard errors.

[22] Due to small sample size (pedestrian and pedacycle accidents account for only 15% of my sample), I am unable to address the question of whether pedestrians, or school-children in particular, would experience an even larger increase in the risk of being hit by a vehicle due to the darkened mornings of DST. Using the same RD design on this limited sample yields imprecise point estimates of similar magnitude to those using the full sample.

2007-2011 data. Running the same RD strategy measures the impact of these transition dates in years where there was no actual shift between Standard Time and DST on these dates. If these dates, rather than DST are responsible for the increased crash counts, this test should reveal a similar increase in crashes to those seen in columns 1-5. The zero result in column 6 suggests that the increase in crashes is not simply due to the transition dates, but due to the actual policy.[23]

To address the concern that my results are driven by how I adjust the crash count for the transition date, I run two additional specifications. First, I follow the method used by Janszky et al. (2012) and multiply the the crash count on the transition date by 24/23rds to calibrate for the shorter time period. Alternatively, I throw out the transition date altogether. In both cases, results are qualitatively identical to my main specification (see Table 2). The remainder of Table 2 shows that results are robust to alternative kernel choice, while Table 3 shows they are robust to using a global polynomial RD design. Overall, these results demonstrate that spring transition into DST is associated with a significant increase in fatal crashes. Now I turn to the fall transition to test whether there is an analogous reduction in crashes when leaving DST.

6.2 Fall RD Design

Figure 5 illustrates the regression discontinuity strategy for the fall. In contrast to the spring, the residual plot looks quite smooth as it crosses the fall transition date. Table 4 presents the corresponding regression results. Just as the residual plot suggests, the preferred specification in column 1 indicates no significant change in fatal crashes associated with leaving DST. This result is robust to alternative bandwidths (columns 2-3) and splitting the sample into just the old October or current November transition date

[23]The negative point estimate would suggest that, if anything, my results understate the true impact of the spring transition into DST.

Table 2: RD estimates of the impact of entering DST on fatal crashes - additional robustness

	Alternative Kernels			24/23rds	No Trans
	(1)	(2)	(3)	(4)	(5)
DST	0.0631**	0.0587*	0.0584*	0.0566*	0.0685**
	(.0309)	(.0314)	(.0312)	(.0307)	(.0340)
Kernel	Uni	Tri	Epa	Uni	Uni
# days left	18	22	20	18	17
# days right	19	23	21	19	17

Dependent Var: Log fatal crashes; all specs use day-of-week and year dummys, a first order polynomial and the bandwidth selector of Calonico, Cattaneo, and Titiunik (2012). DST is the estimate of the discontinuity in fatal crashes that occurs immediately following the spring transition. Uni refers to a uniform kernel; Tri refers to a triangular kernel; Epa refers to an Epanechnikov kernel. 24/23rds is an alternative correction for the spring transition date where the crash count is weighted as 24/23rds. No Trans drops the spring transition date from the sample. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 3: RD estimates of the impact of entering DST on fatal crashes - global polynomial regressions

	Alt Polynomials				Alt Bandwidth	
	(1)	(2)	(3)	(4)	(5)	(6)
DST	0.0805***	0.0844***	0.0646*	0.0727**	0.0828*	0.0583***
	(0.0299)	(0.0302)	(0.0355)	(0.0299)	(0.0434)	(0.0212)
Bandwidth	30	30	30	30	15	60
Polynomial Order	4	3	5	4	4	4
Additional Controls	No	No	No	Yes	No	No

Dependent Var: Log fatal crashes; all specs use day-of-week and year dummys. DST is the estimate of the discontinuity in fatal crashes that occurs immediately following the spring transition. Additional controls consist of national gasoline prices, the S&P 500 index (both in log form) and holiday dummies. Bandwidth is # of days on each side of the transition. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4: RD estimates of the impact of leaving DST on fatal crashes

				2002-2006	2007-2011	Placebo
	(1)	(2)	(3)	(4)	(5)	(6)
Leaving DST	0.0018	0.0226	0.0026	0.0189	-0.0233	0.0231
	(.0247)	(.0207)	(.0175)	(.0331)	(.0542)	(.0236)
Bandwidth	CCT	IK	CV	CCT	CCT	CCT
# days left	18	41	62	13	11	18
# days right	19	42	63	14	12	19

Dependent Var: Log fatal crashes; all specs use day-of-week and year dummys, a first order polynomial and a uniform kernel. Leaving DST is the estimate of the discontinuity in fatal crashes that occurs immediately following the fall transition out of DST. Placebo assigns the current November transition date to 2002-2006 data and the old October transition date to the 2007-2011 data. CCT refers to the bandwidth selector of Calonico, Cattaneo, and Titiunik (2012); IK is Imbens and Kalyanaraman (2012); CV is the cross-validation method of Ludwig and Miller (2007). Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

(columns 4-5). Using an analogous placebo test to that used in the spring suggests that these transition dates do not systematically alter crash risk independent of a policy. Taken as a whole, the transition from DST back to Standard Time does not reduce fatal crash risk in the same way entering DST increases risk. I now turn back to the mechanisms through which DST could impact crash risk to explain this asymmetric effect.

6.3 Mechanisms

The spring transition is subject to both the light and sleep mechanism. Hence, the 6.3% increase in fatal crashes could be partially due to each mechanism. The most parsimonious method for decomposing this result into each mechanism uses only aggregate results from the spring and fall. Given the fall transition is not subject to any change in sleep quantity, it isolates the light mechanism. The aggregate effect of zero when leaving DST in the fall suggests no net impact of DST through the light mechanism. Differencing the spring estimate of 6.3% (light and sleep mechanism active) and the fall estimate

of zero (only light mechanism active) provides suggestive evidence that the impact of the spring transition should be attributed solely to the sleep loss mechanism.[24] However, differences in sunrise and sunset times and the potential for differences in driver behavior between the spring and fall transitions prevent this from being an ideal comparison. To further disentangle the mechanisms, I use the initial RD framework with sub-samples of hours selected to isolate the impact of one mechanism or the other.

6.3.1 Light

Since only the light mechanism is active during the fall, the aggregate fall effect of zero suggests no net impact through this channel. To determine if light has become altogether unimportant as a fatal crash risk factor, perhaps through improved vehicle lights, I further explore the light mechanism by examining sub-samples of hours closest to sunrise and sunset.[25] Upon leaving DST in the fall, an hour of light is removed from the evening and returned to the morning. If light remains an important fatal crash risk factor, additional morning light should create a safer atmosphere for driving during morning hours. Likewise, removing light from the evening should create a more dangerous driving atmosphere during this time. To test this hypothesis, I break the sample into a set of morning hours (4-9 a.m.) and evening hours (3-8 p.m.). Then I run the initial RD analysis on these subsamples for the fall transition. If light remains an important factor in fatal crash risk, leaving DST should lead to fewer morning crashes

[24] The aggregate estimates for leaving DST tend to be positive (though insignificant). By symmetry, if leaving DST increases fatal crash risk this implies that entering DST reduces fatal crash risk. Hence, if anything, the light mechanism reduces crashes during DST (as suggested by Broughton, Hazelton, and Stone (1999) and Ferguson et al. (1995)). As such, the 6.3% increase in crashes in the spring is, if anything, a downwardly biased estimate of the sleep mechanism.

[25] Since 2003 BMW, Toyota and others have released vehicles with Adaptive Front-Lighting Systems (AFS). AFS are designed to optimize headlight direction and volume in response to steering, ambient weather, visibility conditions and speed.

Table 5: RD estimates of the influence of ambient light on fatal crashes when leaving DST

	All Hours	Morning			Evening		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Leaving DST	0.0018	-0.1631**	-0.1182**	-0.1482**	0.1208**	0.2093***	0.1614***
	(.0247)	(.0703)	(.0555)	(.0657)	(.0506)	(.0499)	(.0323)
Bandwidth	CCT	CCT	IK	CV	CCT	IK	CV
# days left	18	16	30	57	13	61	16
# days right	19	17	31	58	14	60	17

Dependent Var: Log fatal crashes; all specs use day-of-week and year dummys, a first order polynomial and a uniform kernel. Leaving DST is the estimate of the discontinuity in fatal crashes that occurs immediately following the fall transition out of DST. "Morning" refers to 4-9am; "Evening" is 3-8pm. CCT refers to the bandwidth selector of Calonico, Cattaneo, and Titiunik (2012); IK is Imbens and Kalyanaraman (2012); CV is the cross-validation method of Ludwig and Miller (2007). Robust standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

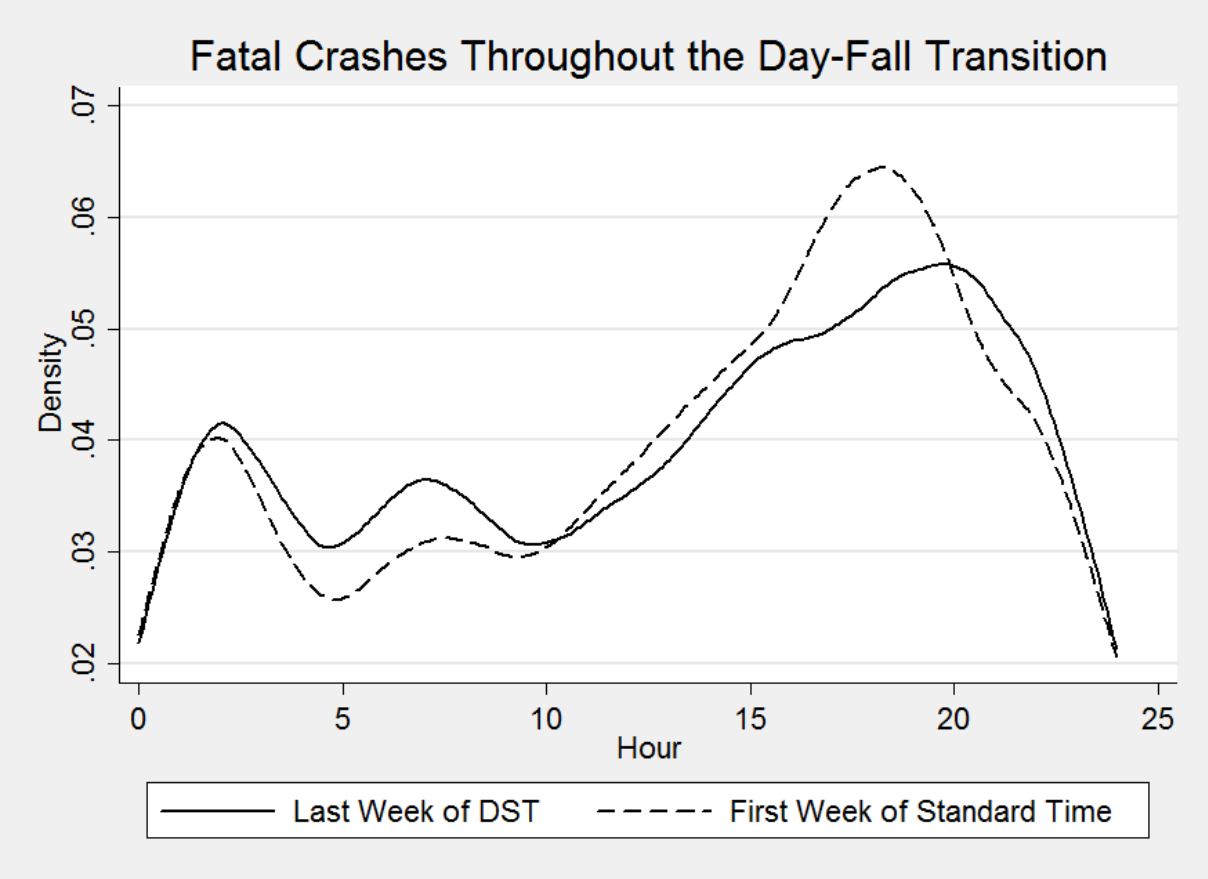
(more light) and additional evening crashes (less light). If no change in crashes is seen, it is likely that light no longer plays an important role in fatal crashes. Table 5 details the results.

Across different bandwidths, leaving DST is associated with a significant reduction in fatal crashes during the morning (more ambient light). Conversely, evening hours (less ambient light) are always associated with a significant increase in fatal crashes. These results suggest that light still plays an important role in fatal crash risk. However, the aggregate zero effect (Column 1) suggests these impacts balance out and light has no net impact through DST. Crashes are simply reallocated between the morning and the evening. This reallocation can be seen more clearly in the kernel density function in Figure 7.

6.3.2 Sleep

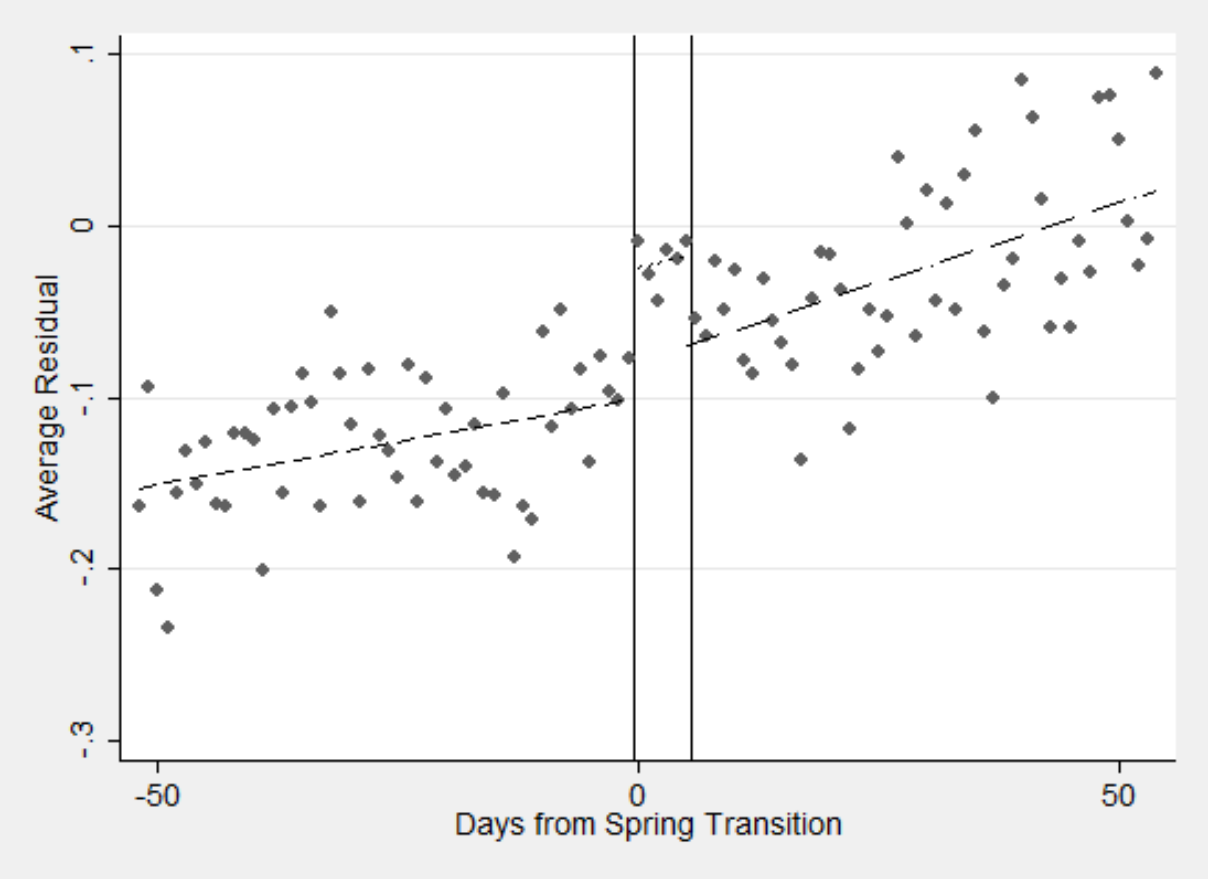
The spring transition is subject to both the sleep and light mechanisms. However, my estimates for the fall transition suggest that the net impact of the light mechanism

Figure 7: Reallocation of Fatal Crashes (Fall Transition)



Notes: The kernel density functions use an Epanechnikov kernel. First week of standard time begins on the 25-hour transition date (Sunday).

Figure 8: Spring Residual Plot – Six Day Sleep Impact



Notes: The residuals are generated from a regression of ln(fatal crash count) on day-of-week and year dummies. Each point is the average of all residuals for that date relative to the spring transition. Fitted lines impose linear seasonal trend on residuals.

is zero. Taking a closer look at the spring residual plot in Figure 8 provides a clearer picture of what is occurring right at the spring transition. There is a discontinuous jump in fatal crashes that seems to persist for the first six days of DST, before jumping back down to essentially the same seasonal path seen during Standard Time. Since the light mechanism is in effect for the entire period of DST, this data pattern is inconsistent with a light impact – we would not expect the crash count to jump back down. However, a shock to sleep should only be felt in the initial period following the transition, before dissipating – exactly the phenomenon seen here.

Table 6: RD estimates of the influence of sleep loss on fatal crashes

	All Hours	Least Light Impacted Hours		
	(1)	(2)	(3)	(4)
DST	0.0631**	0.0484	0.0601**	0.0773***
	(.0309)	(.0360)	(.0250)	(.0258)
Bandwidth	CCT	CCT	IK	CV
# days left	18	17	36	57
# days right	19	18	37	58

Dependent Var: Log fatal crashes; all specs use day-of-week and year dummys, a first order polynomial and a uniform kernel. DST is the estimate of the discontinuity in fatal crashes that occurs immediately following the spring transition. Least Light Impacted Hours are 9am-3pm and 8pm-4am. CCT refers to the bandwidth selector of Calonico, Cattaneo, and Titiunik (2012); IK is Imbens and Kalyanaraman (2012); CV is the cross-validation method of Ludwig and Miller (2007). Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

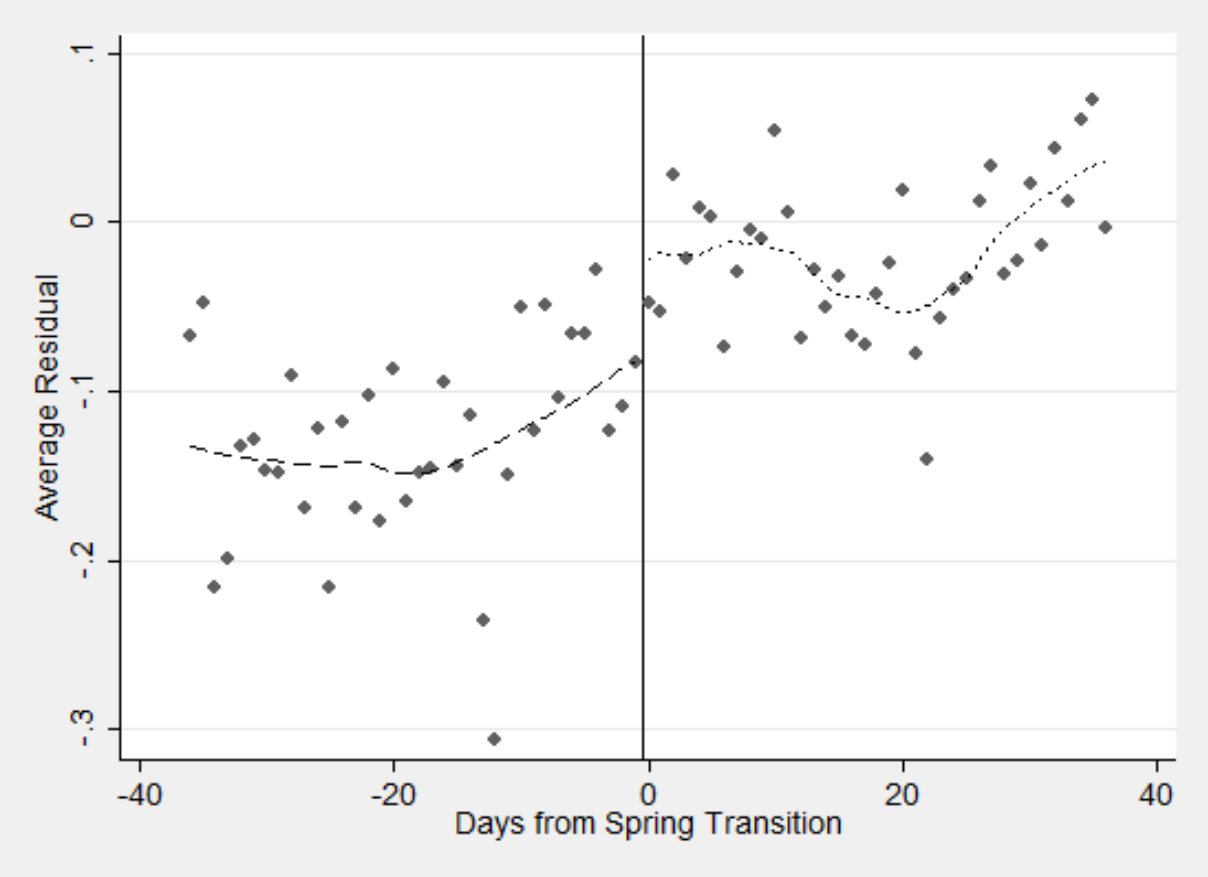
To pry further at the sleep mechanism, I focus on a sub-sample of hours furthest away from sunset and sunrise to mitigate the light impact.[26] Figure 9 illustrates the discontinuity while Table 6 provides the regression results. The point estimates are quite similar to the full day impacts and are significant using two of the three bandwidth selectors. This suggests that it is the sleep mechanism, not light, that causes the short-run increase in fatal crashes following the spring transition. To further investigate the mechanisms and to determine the length of this sleep impact, I turn to the fixed effects model.

6.4 Fixed Effects Model

Table 7 presents the results from the FE model. The point estimates represent the average impact of DST over the full range of switching dates (dates that are DST in some years and Standard Time in others), rather than just at the threshold. While the

[26] I say "mitigate" not "eliminate" because the angle of the sun and moon are still altered even in these hours of full light and full darkness.

Figure 9: Spring Residual Plot – Least Light Impacted Hours



Notes: The residuals are generated from a regression of ln(fatal crash count) on day-of-week and year dummies. Each point is the average of all residuals for that date relative to the spring transition. Fitted lines are results of locally weighted regression. Least light impacted hours are 9am-3pm and 8pm-4am.

initial columns examine the spring DST period as a whole, columns 3-7 break spring DST down into three components (i) the first six days of DST, where the sleep effect should be felt most strongly;[27] (ii) the next eight days of DST, the longest any sleep study suggests a sleep impact could persist; and (iii) the remainder of spring DST with common support, days in which only the light mechanism should remain present.

[27] I choose six days based on the appearance of the residual plot seen in Figure 8. This covers the Sunday-Friday following the spring transition and is consistent with the literature on how long DST impacts sleeping patterns.

Table 7: FE estimates of the impact of DST on fatal crashes

	All Hours				Least Light Affected	Morning	Evening
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Spring DST	0.0340**	0.0335**					
	(0.0164)	(0.0165)					
First 6 Days of DST			0.0559**	0.0570**	0.0481*	0.206***	0.00745
			(0.0231)	(0.0233)	(0.0279)	(0.0525)	(0.0396)
Next 8 days of DST			0.0285	0.0279	0.0385	0.134**	-0.0668
			(0.0200)	(0.0201)	(0.0254)	(0.0539)	(0.0417)
Remainder of Spring DST			0.0181	0.0161	0.00764	0.121**	-0.0260
			(0.0196)	(0.0197)	(0.0231)	(0.0521)	(0.0389)
Fall DST	0.0280	0.0272	0.0272	0.0263	0.0541*	0.225***	-0.120***
	(0.0246)	(0.0245)	(0.0246)	(0.0246)	(0.0320)	(0.0653)	(0.0448)
Additional Controls	No	Yes	No	Yes	Yes	Yes	Yes
Observations	3,652	3,652	3,652	3,652	3,652	3,652	3,652
R-squared	0.755	0.759	0.755	0.760	0.770	0.295	0.421

Dependent Var: Log fatal crashes; all specs use day-of-year, day-of-week and year dummys. Remainder of Spring DST is an indicator variable equal to one if the day occurs after the first two weeks of DST and by April 7th, the final spring switching date. Fall DST is an indicator variable equal to one if the day falls under DST and occurs on Oct 25th or later, the first fall switching date. Additional Controls are ln(gas prices), ln(S&P index) and dummies for nonstationary holidays. Morning refers to 4-9am; Evening refers to 3-8pm; Least Light Affected are the remaining hours. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Beginning with the entire spring period, column 1 shows that spring DST is associated with a significant 3.4% increase in fatal crashes over the roughly one month of switching dates. The fall estimate is insignificant from zero, again suggesting no impact of DST in the fall.[28] In addition to day-of-year fixed effects, column 1 uses just day-of-week and year dummies, the same controls used in the RD design. Column 2 includes additional covariates for nonstationary holidays, gasoline prices and the value of the S&P 500 index.[29] Results are quite stable across columns and continue to suggest that DST causes a significant increase in crashes during the spring and has no effect during the fall.

Turning to columns 3-4, the results are broadly consistent with a sleep impact that diminishes further from the spring transition and no net impact from reallocating light. The first six days of DST experience a significant 5.6% increase in fatal crashes, quite similar to the 6.3% increase found in the RD design. The point estimate shrinks to an insignificant 2.9% during the next eight days and diminishes further to 1.8% for the remainder of the spring. During both time periods in which only the light mechanism is active, the fall and the spring following the first two weeks, there is no significant change in crash counts. Including additional controls in column 4 to help proxy for the character and amount of vehicle miles traveled leaves results qualitatively identical.

Columns 5-7 explore these impacts across different times of day, reinforcing previous findings regarding the sleep mechanism. Column 5, uses just the subsample of hours least effected by the light mechanism, effectively isolating the sleep mechanism. The 4.8% increase in crashes during the first six days of DST provides a measure of the impact of just the sleep deprivation mechanism on crashes during these hours. Across each subsample of hours, the point estimates drop from the first six days of DST to

[28] The fall estimates are less precise because there was only a 1-week extension to DST in the fall, providing fewer switching dates than in the spring.

[29] Adding each additional covariate individually leaves results qualitatively identical.

beyond the first two weeks of DST in the spring. This suggests that across all hours, mitigating the sleep mechanism reduces fatal crash risk. Overall, the body of evidence from the FE model aligns with that found from the RD model. There is a significant short-term increase in fatal crashes following the spring transition, consistent with a detrimental impact of sleep loss. Now I turn to plausible alternative explanations for this short-term spike in fatal crashes.

6.5 Alternative Explanations

A key omitted variable in this analysis and previous studies is Vehicle Miles Traveled (VMT). If VMT increases at the DST transition date, this behavioral change could be driving results rather than sleep loss. While national VMT data is not available, the Performance Measurement System (PeMS) in California tracks VMT on many major highways within the state. Using the same regression discontinuity model from equation 1 with log(VMT) as the dependent variable yields an insignificant 0.016% increase in VMT. To the extent that driving habits on these California roadways are representative of national driving patterns, this suggests VMT is not the cause of increased crashes.

Adverse weather conditions increase the risk of fatal crashes (Fridstrom et al., 1995). Although weather is pseudo-random, if adverse weather occurred just following the spring transition, this could lead to the short-term increase in fatal crashes. Using a FARS variable that indicates weather conditions at each fatal crash, I create a variable for the ratio of crashes within a day that are impacted by weather. Using the regression discontinuity model from equation 1 with weather-ratio as the dependent variable I find an insignificant 1.2 percentage point decrease in weather related crashes.[30]

[30] The residual plots and regression output for both of these "alternative explanations" are available in the appendix.

This analysis suggests that some of the most likely alternative pathways cannot explain the increase in fatal crashes. Further, if the increase is due to adjusting to a new schedule, the same increase should occur immediately following the fall transition, a phenomenon that we do not see. While this is not an exhaustive list of competing explanations, the balance of evidence points strongly towards DST increasing fatal crash risk, through the mechanism of sleep deprivation. In the next section, I explore whether this result varies by region.

6.6 Geographical Heterogeneity

At the national level, the spring transition into DST leads to a significant increase in fatal crashes. However, this could be due to a constant treatment effect where all regions experience the same 6% increase in crashes, or a heterogeneous treatment effect where some regions experience a larger increase and others experience little or no effect. In this section, I explore two pathways through which geography could lead to heterogeneous impacts of DST, one through the sleep mechanism and the other through the light mechanism.

Sleep deprivation could be more detrimental when driving in already dangerous area. If there are more situations where a delayed response can lead to a crash, the sleep mechanism has more scope to operate. To test this hypothesis, I split my sample in two based on the median number of fatal crashes per capita in each county.[31] The counties with a higher per capita fatal crash rate, I refer to as high risk counties. Running the RD analysis with these subsamples (Table 8) provides weak evidence that high risk counties are subject to a larger initial increase in fatal crashes (in percentage terms) than their low risk counterparts. While the estimates may not be statistically different

[31] 2010 census counts used for county population.

Table 8: RD estimates of the impact of entering DST on fatal crashes, by county risk level

	High Risk Counties			Low Risk Counties		
	(1)	(2)	(3)	(4)	(5)	(6)
DST	0.0817	0.0919**	0.1213***	0.0466	0.0352	0.0576**
	(.0530)	(.0417)	(.0421)	(.0346)	(.0222)	(.0237)
Bandwidth	CCT	IK	CV	CCT	IK	CV
# days left	23	50	57	16	42	57
# days right	24	51	58	17	43	58

Dependent Var: Log fatal crashes; all specs use day-of-week and year dummys, a first order polynomial and a uniform kernel. DST is the estimate of the discontinuity in fatal crashes that occurs immediately following the spring transition. High and Low Risk Counties are based on a cut at the median county of fatal crashes per capita based on 2010 county population. CCT refers to the bandwidth selector of Calonico, Cattaneo, and Titiunik (2012); IK is Imbens and Kalyanaraman (2012); CV is the cross-validation method of Ludwig and Miller (2007). Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

at conventional levels, in all cases the point estimate for high risk counties is above that of low risk counties. This provides suggestive evidence that sleep loss is more detrimental when performing a more difficult task.

If ambient light is more important in certain hours than others, heterogeneity in sunrise and sunset times within a time zone could lead to differential impacts of DST. Sunrise occurs earliest in the Eastern portion of any time zone; in Boston, sunrise the day before DST occurs at 6:07 a.m. whereas in Louisville, Kentucky, it occurs at 7:04 a.m. In Boston, the onset of DST moves sunrise back an hour to roughly 7 a.m. while in Louisville sunrise is moved to roughly 8 a.m. If light is more important for fatal crashes (perhaps due to more driving) during the 7-8 a.m. hour relative to the 6-7 a.m. hour, Louisville should experience a bigger morning increase in fatal crashes (in percentage terms) than Boston.[32] To test this mechanism, I split the sample into an

[32]In the evening, sunset shifts from 17:45 to 18:45 in Boston and 18:45 to 19:45 in Louisville. Again, it would appear that Boston is helped more, as 17:45 to 18:45 is more of a peak travel time than 18:45-19:45.

Table 9: RD estimates of the impact of entering DST on fatal crashes - geographical impacts

	Eastern Portion of Time Zone			Western Portion of Timezone		
	(1)	(2)	(3)	(4)	(5)	(6)
DST	0.0737	0.0621	0.784**	0.1066***	0.0525*	0.0726**
	(.0502)	(.0386)	(.0391)	(.0343)	(.0308)	(.0299)
Bandwidth	CCT	IK	CV	CCT	IK	CV
# days left	21	42	57	23	48	57
# days right	22	43	58	24	49	58

Dependent Var: Log fatal crashes; all specs use day-of-week and year dummys, a first order polynomial and a uniform kernel. DST is the estimate of the discontinuity in fatal crashes that occurs immediately following the spring transition. The Eastern Portion of a TZ are the roughly 1/3 of crashes most Eastern based on latitude within a timezone, the Western Portion the same for the West. CCT refers to the bandwidth selector of Calonico, Cattaneo, and Titiunik (2012); IK is Imbens and Kalyanaraman (2012); CV is the cross-validation method of Ludwig and Miller (2007). Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Eastern, Western, and Central third of each timezone.[33]

Table 9 shows the RD results. In contrast to what might be expected based on common commute times, results are quite similar for both areas. Figure 2 helps to elucidate this finding. While the darkened hour in the Eastern portion of time zones has fewer fatal crashes and the brightened hour has more fatal crashes, it is a very minor difference. Further, the average sunset and sunrise times in the Eastern and Western portion of a timezone is closer than the full hour seen in the Boston - Louisville example. This geographic heterogeneity could be explored further in other applications where higher frequency events would increase the power of the test and allow for more narrow geographic areas than one third of a timezone.

[33] I split each timezone into East, West, and Central thirds based on number of fatal crashes in each portion (rather than by population or landmass).

7 Conclusion

Daylight Saving Time is one of the most practiced policies across the globe, impacting over 1.5 billion people. Despite this worldwide coverage, many of the impacts of DST remain empirical questions. I exploit the discrete nature of transitions between Standard Time and DST, and variation in the coverage of DST created primarily by a 2007 policy change, to estimate the impact of DST on fatal vehicle crashes. My main finding is that the spring transition into DST increases fatal crash risk by 5.4-7.6%.

I employ three tests to determine whether this result is due to shifting of ambient light or sleep deprivation caused by the 23-hour transition date. These tests reveal that while ambient light reallocates risk within a day, it does not contribute to the increase in crashes. All three tests suggest that the sleep deprivation is driving the increase in fatal crashes. Consistent with literature investigating the impact of DST transitions on sleep, the impact persists for the first six days of DST. Back of the envelope calculations suggest that over the ten year study period, DST caused 302 deaths at a social cost of $2.75 billion.[34]

In terms of DST, this result should be viewed as one piece of the puzzle, to be examined in conjunction with research on other impacts of DST. In previous research, when a benefit of DST is found it tends to be through the light mechanism. More light in the evening has benefits at reducing crime (Doleac and Sanders, 2013) and encouraging exercise (Wolff and Makino, 2013).[35] When costs are found, similar to my study, it tends to be due to sleep loss or disruptions associated with transitions (Janszky et al., 2012).

[34]Social cost is calculated as follows: Multiplying the 5.6% increase found in the FE model by the 489.3 fatal crashes averaged on Sundays-Fridays in March and April yields 27.4 additional fatal crashes per year. Multiplying this by the 1.104 fatalities per crash observed over my sample and the 10 year study period yields and extra 302 deaths over 10 years. Applying the Department of Transportation's $9.1 million value of a statistical life, this a $2.75 billion social cost.

[35]One concern about DST is that morning rise time relative to sunrise time is an important factor in clinical depression (Olders, 2003).

Taking these points in combination, an ideal policy solution would leave the benefits of DST intact while eliminating the damage caused by the spring transition. Before a significant policy change is made, further research should be conducted on the welfare effects of the policy.

Finally, this paper fits into the small but growing literature examining the impact of sleep on worker productivity (Kamstra, Kramer, and Levi, 2000; Lockley et al., 2007; Barnes and Wagner, 2009; Wagner et al., 2012). Although fatal vehicle crashes are an extreme measure of productivity, driving is an activity that over 90% of American workers engage in (Winston, 2013) and DST provides an exogenous shock to sleep quantity. The increased risk of a fatal vehicle crash suggests significant costs of sleep deprivation, even when undertaking a routine task. Given the ongoing trend towards less sleep, particularly among full-time workers (Knutson et al., 2010), it is important that researchers continue to investigate the relationship between sleep and productivity. My results represent a lower bound for the overall cost of DST through sleep deprivation, since reductions in workplace productivity are unaccounted for.

Chapter II
Alternative Compensation and Performance: Evidence From Staking in Online Poker

8 Introduction

A central tenet of personnel economics is that workers respond to incentives. Given that monitoring costs in most work settings can be prohibitively high, workers may shirk duties in the absence of incentives to align their interests with those of the firm (Nagin et al., 2002). In response to this, there has been renewed interest in the use of performance based contracts – contracts in which workers are paid based on output. Some form of performance pay was used in 39% of US private sector jobs during 2013 (Gittleman and Pierce, 2013). However, empirical studies on performance pay are relatively scarce.

Existing evidence on the efficacy of performance pay is primarily clustered on two types of workers. The first set of studies focus on rote jobs where output is easily measurable. Here, there is a consensus that performance pay leads to significant increase in worker productivity.[36] The second set of studies examine the impact of CEO compensation on firm performance.[37] While there is less consensus in this area, recent evidence suggests that risk-aversion on the part of the CEO can lead to worse firm outcomes when CEO pay is tied more closely to stock returns (Brick, Palmon, and Wald, 2012). Less is known about the impact of performance pay on those in the middle to high end of

[36]Lazear (2000), uses data from windshield installer Safelite to demonstrate that individuals are more productive when paid via a piece-rate than an hourly wage. Selection effects, induced by the piece-rate scheme, further improve the quality of the average worker. Similarly, Paarsch and Shearer (1999) find that tree planters produce higher levels of output when facing a piece-rate.

[37]See Tosi et al. (2000) for a summary of these studies.

the income distribution, particularly regarding workers engaged in primarily cognitive tasks.[38] We focus on this group, using a unique panel dataset that tracks workers as they are able to adjust how they are paid for performance. We find that performance is best when payment is made solely based on outcomes, with performance declining as more weight is given to a fixed component of pay.

The analysis in this paper is based on data from over 100,000 online poker tournaments played by 98 players from May 2009 through December 2010.[39] The advent of a new marketplace allowed players to seek "staking" for individual poker tournaments, an arrangement in which investors pay players a fixed fee for participating in a tournament in return for a set percentage of prize money. Unstaked, a player's pay for a given poker tournament is based solely on their finish position in that tournament. When staked, a player's pay is based on two components: i) a fixed fee received from the investor, and ii) the residual share (after the investor receives their share) of any prize earned for performance in the poker tournament.

Staking likely alters performance in two competing ways. First, staking creates the alternative compensation structure described above, which reduces the players marginal return to an improvement in rank. Reducing the marginal return to finish position in a tournament structure has been shown to reduce the effort level of competitors (Ehrenberg and Bognanno, 1990a,b). Second, staking loosens the financial constraints facing players.[40] This could allow players, who face a choice among risky alternatives, to adopt a higher-risk strategy with a higher expected value.

[38] Fernie and Metcalf (1999) examine professional horse jockeys, a group that is high-skill and high-income, but their task is largely physical. They find that jockeys perform better when their pay is tied to the race prize than when receiving a fixed fee.

[39] Levitt and Miles (2014) demonstrate that poker tournaments are skilled based. Using data from World Series of Poker tournaments, they find that *a priori* identified professional players earn a return on investment that is 46 percentage points higher than amateurs.

[40] Players must pay an entry fee to participate in each tournament.

Exploiting variation in staking status within a player, we estimate the impact of staking on player performance. We adopt a player by skill-tier fixed effects approach, comparing results for a given player at a given entry fee level when staked, to results for the same player at the same entry fee level when unstaked. Across different measures of performance, we find strong evidence that performance is worse when staked, suggesting that the muting of incentives created by staking is the dominant factor influencing performance.

Despite the within-player comparison group used for identification, staking is not randomly assigned. A player selects into staking, posting an advertisement only for those tournaments for which they desire staking. As such, selection bias could play a role in our results – a player may only seek staking when they possess private information that makes staking advantageous for them. This would bias our results towards our current finding of worse performance when staked. To address this concern, we perform three tests. First, we look for unusually strong results before the first staking incident. This could indicate that a player appears more skilled to investors than they truly are, and that the player is trying to capitalize on this appearance by garnering a higher fixed fee. Worse results when staked could be regression to the mean. Second, we use variation in the amount that players are staked (what percent of prize money they have sold) only within staked tournaments. Finally, we analyze the variation in percent staked created on the investor side of the market (situations in which a player does not receive the full amount of staking requested). While selection likely plays a small role, all three tests provide evidence that the alternative compensation structure created by staking is the main driver of worse performance.

Our results suggest that the muting of incentives caused by staking reduces overall performance of the worker. In our setting, where an unstaked player is operating like a firm (bearing all risk and receiving all profits), this implies that moving from a sit-

uation where the worker is the firm to one where incentives are closely aligned, leads to a significant productivity loss. We acknowledge that part of this loss is potentially due to selection, which has been shown to play a large role in sorting into incentive schemes across workers (Lazear, 2000; Dohmen and Falk, 2011). However, our empirical tests suggest that this is primarily an incentive story. Overall, our results confirm the prediction of tournament theory that larger prizes induce higher effort levels from competitors. This has implications for firms, where promotions often follow a tournament structure with employees promoted based on their performance relative to other employees (Lazear, 1992; Baker, Gibbs, and Holmstrom, 1993, 1994a,b; Bognanno, 2001). Our results suggest that increasing the prize (the value of the promotion) can be an effective way to increase productivity, even when output is highly variant as is the case in poker tournaments.

The remainder of the paper is organized as follows. Section 2 briefly describes the features of online poker tournaments and the market for staking that are crucial for understanding our empirical analysis. Section 3 describes the data, while section 4 outlines our empirical framework and central estimation equations. Section 5 presents the results of our analysis, including tests that differentiate between potential mechanisms. Section 6 concludes with a discussion of the implications of our findings.

9 Online Poker Tournaments and the Market for Staking

9.1 Online Poker Tournaments

The typical online poker tournament is open to any individual willing to pay the entry fee (referred to as *buyin* hence forth). In exchange for the buyin, participants receive a predetermined amount of tournament chips. Players are randomly assigned a table and the tournament plays out continuously, with participants being eliminated when they run out of chips, until only one player remains (with all of the chips). Prizes are awarded based on inverse order of elimination, with approximately 10% of the field receiving payouts. The winner receives the largest share of the prize pool, followed by the last player eliminated and so on.

The prize pool is funded by the buyins of all competitors, with some portion of the fee going to the hosting site (about 8% is typical in our sample). Prizes increase nonlinearly with finish position and follow the general structure seen in Figure 10. Notably, the marginal return from moving up one finish position from 3rd to 2nd is worth 3.4% of the prize pool whereas moving up from 19th to 18th is worth only 0.07%.[41] Across tournaments the structure of prizes based on finish percentile is virtually identical, although the level of prize money varies across tournaments based on the buyin and number of participants. The top heavy prize structure and the stochastic component of poker create a high level of variance in the earnings of tournament players.[42] To reduce variance, some poker players turn to a secondary market to reduce their risk.[43]

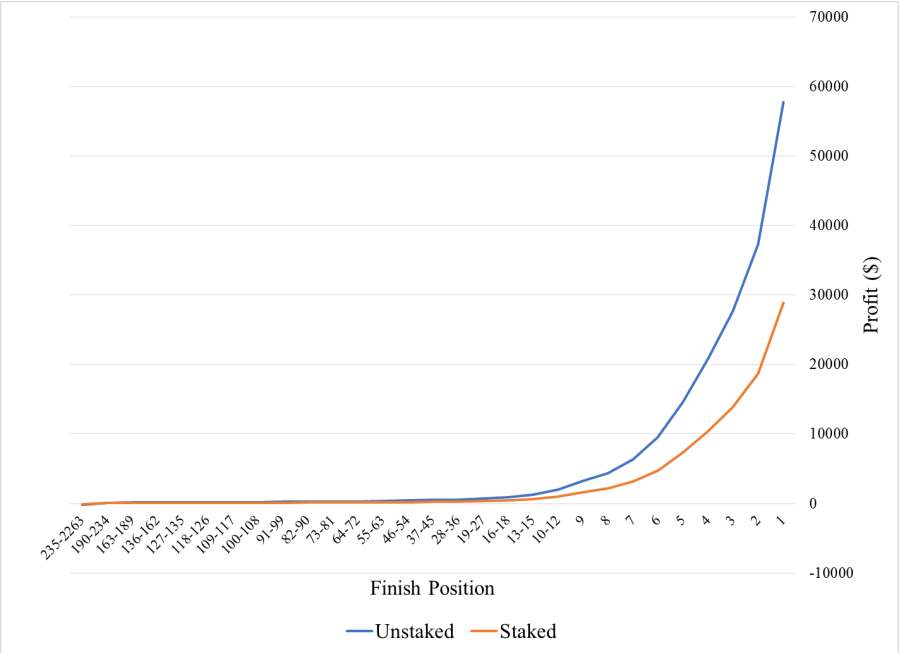
[41] Marginal returns are based on a field size of 2263, the mean staked tournament in our sample.

[42] See (Levitt, Miles, and Rosenfield, 2012) for discussion of the relative importance of skill versus luck in poker.

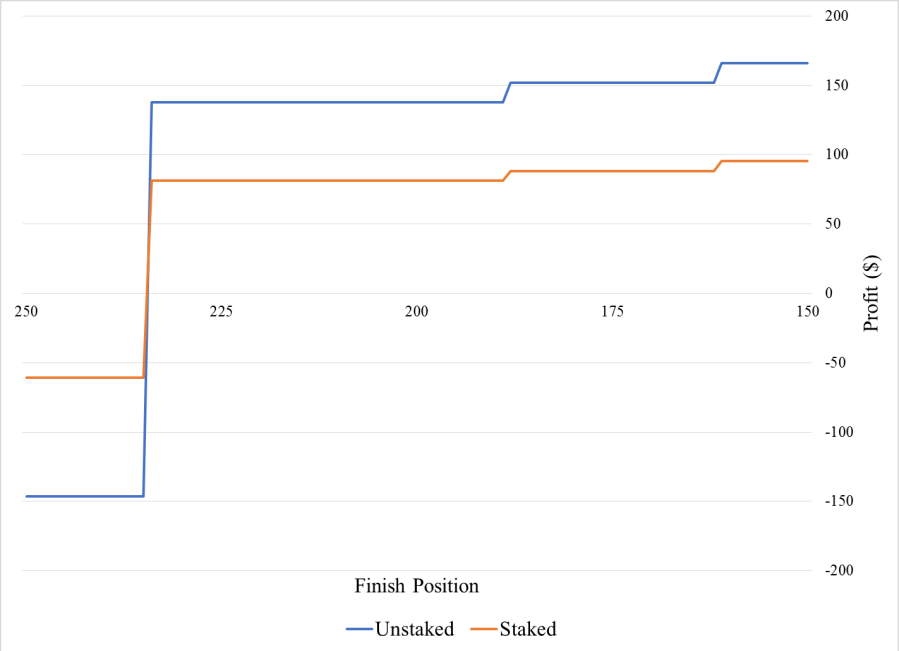
[43] Many players specify in their advertisements why they are seeking staking, with reducing the variance of outcomes being the most common stated reason.

Figure 10: The Impact of Staking on Poker Tournament Prizes

(a) Full Distribution of Prizes



(b) At First Prize Level



Notes: Figure 1a depicts the distribution of profit for the mean staked tournament in our sample, under the conditions of no staking and staking of 50% with average markup. Figure 1b zooms in on finish positions near the first prize.

9.2 The Market for Staking

Staking is an arrangement in which an investor(s) pays a portion of a player's buyin for a specific poker tournament, in return for an agreed upon percentage of any prize money won by that player in that tournament. While informal staking arrangements have likely existed since the advent of poker, formal marketplaces are a relatively new phenomenon.[44] Our data come from the staking marketplace on twoplustwo.com (the largest poker strategy forum on the internet). Here, the market generally proceeds in three stages: (i) the player advertises the tournament(s) for which they are seeking staking and the terms of the deal; (ii) investors express their intention to purchase some or all of the available stake and send money to the player; and (iii) the player participates in the agreed upon tournament(s) and sends the investor their share of prize money.

Figure 11 walks through this process for a typical example. The advertisement includes the tournament(s) for which staking is being requested and the total amount requested. Often, shares are sold with markup, meaning that an investor must pay more than 1% of the buyin to be entitled to 1% of the prize money. Markup of 15%, a typical amount in our sample, means that an investor must pay 1.15% of the buyin to be entitled to 1% of the prize money. Finally, advertisements provide evidence of previous success, linking to a complete history of all tournaments previously played by the player on the major online poker sites.

Once an advertisement has been posted, any member of the marketplace may post to purchase some (or all) of the stake. As seen in Figure 11, it is common for investors to purchase only a portion of the total amount for sale. Hence, to sell out, a player often receives staking from multiple investors.[45] Once the sale is complete (or the tournament

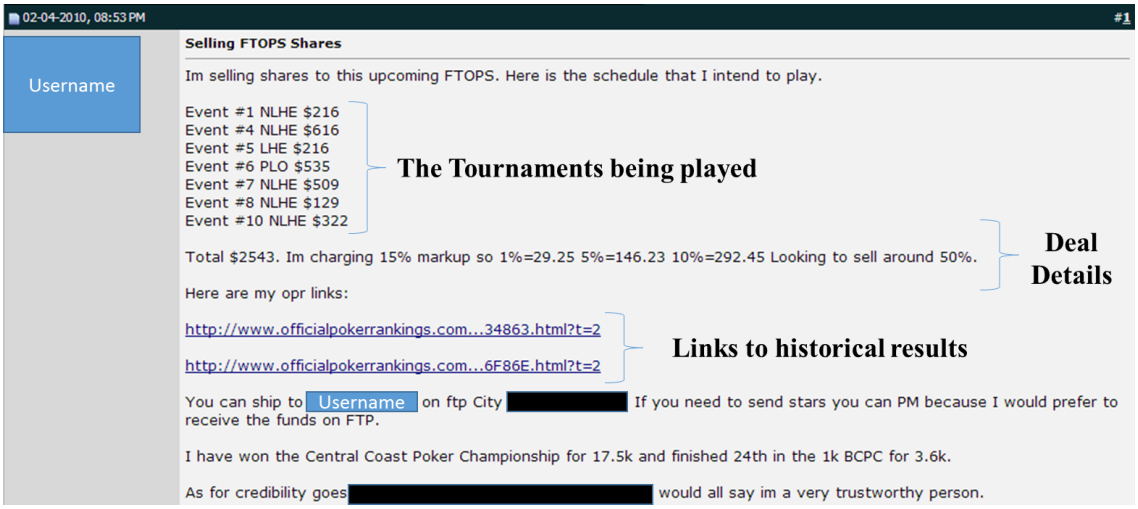
[44] The staking marketplace on twoplustwo.com opened in 2008.

[45] This system creates situations in which a player recieves only some, but not all of their requested level of staking. We use this variation, generated on the investor side of the market, to try to differentiate between mechanisms.

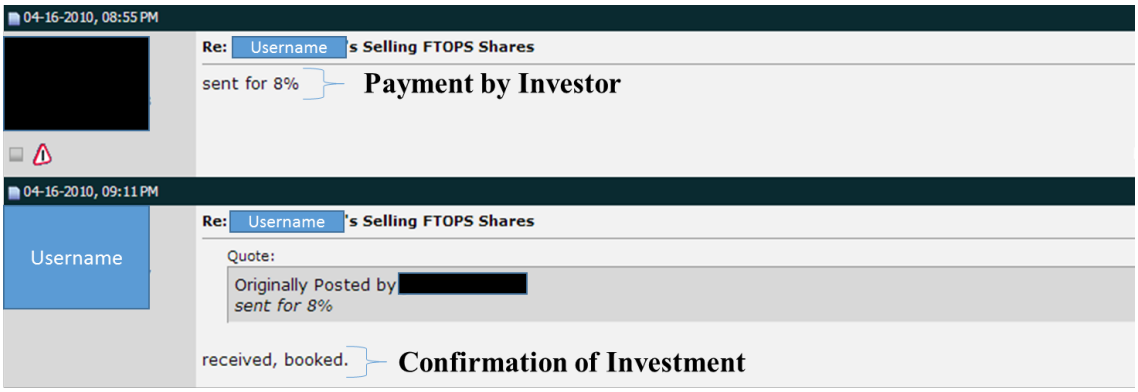
is about to start if the stake does not sell out), the player confirms the receipt of all investor funds. Upon completion of the tournament(s), the player sends the appropriate percentage of prize money to each investor.

Figure 11: Typical Staking Transaction

(a) Phase 1: Advertisement



(b) Phase 2: Investment



(c) Phase 3: Payout

Notes: Staking data come from the marketplace forum on twoplustwo.com. This example, which follows the typical structure of a staking transaction, comes directly from our sample.

While staking can take various forms, the transactions in our sample are all one-off arrangements. The player does not owe investors anything if they do not earn a prize in the staked tournament(s) and the player is not obligated to seek staking again or to give current investors any preference in future sales. Once the stake has been settled, the relationship between player and investor is effectively over.

10 Data

Our staking data come from the staking marketplace on twoplustwo.com. We recorded every transaction occurring from August 2009 through May 2010 for tournaments played on one online poker site, Full Tilt Poker (FTP).[46] To increase the number of observations we extended coverage for staked players backwards, using archived posts, to the first staking incident which occurred in May 2009, and forward through the end of 2010.[47] After cross-checking the player usernames with the other major staking marketplace, Part Time Poker, and dropping players who received staking for unknown tournaments, we were left with 98 players and over 3000 staked tournaments.

We merged this staking dataset with tournament results for these players, gathered from OfficialPokerRankings.com. The tournament results are comprehensive, with one record for every tournament played on FTP for each player in our sample. Each record includes buyin, number of entrants, finish position, prize won, and tournament characteristics. Over the 20-month period, from May 2009 through Dec 2010, our 98-player sample played a total of 100,641 tournaments.[48] Of these, 3,344 were successfully

[46] We choose Full Tilt Poker because it was the largest online site during this timeframe for which a complete history of tournament finishes by player is available.

[47] Online poker in the United States was shut down on Friday, April 15th 2011. We choose to end our sample at the end of 2010 as rumors about the solvency of FTP began in early 2011 and it became common for a dollar on FTP to be sold for less than $1.

[48] A small subset of tournaments (qualifiers and sit-n-gos) were dropped from the sample because there were exceedingly few staked observations and the payout structure differs markedly from the standard structure seen in Figure 10.

matched as staked tournaments. The remaining 97,297 without corresponding staking records are assumed to be unstaked.[49]

Table 10 provides the summary statistics for our sample. Unconditionally, performance in staked tournaments is significantly worse. This is seen in a worse finish percentile, lower return on investment (*buyins won*), and reduced likelihood of having a big or very big win (*win3 and win10* respectively). However, at least some of this performance gap can likely be attributed to the different characteristics of staked tournaments relative to unstaked tournaments. On average, staked tournaments have significantly higher buyins, $157 to $81, and larger field sizes. In the following section, we outline our empirical strategy for estimating the causal impact of staking on performance.

[49]To the extent that we may be missattribute a tournament that was staked privately (not on a marketplace) as unstaked, our estimates represent a lower bound for the impact of staking on performance.

Table 10: Summary Statistics

Variable	Full Sample	Unstaked	Staked	t-Statistic
staked	0.033	-	-	-
	(0.179)			
buyin	83.188	80.667	156.547	**-20.461**
	(113.700)	(107.769)	(213.516)	
entries	1454.040	1426.226	2263.334	**-14.342**
	(2895.252)	(2875.002)	(3332.883)	
prize	0.134	0.133	0.138	-0.782
	(0.340)	(0.340)	(0.345)	
finishpercentile	58.318	58.352	57.310	**2.324**
	(24.780)	(24.753)	(25.522)	
smallwin	0.085	0.085	0.100	**-2.836**
	(0.279)	(0.278)	(0.299)	
win3	0.049	0.049	0.039	**3.051**
	(0.216)	(0.216)	(0.193)	
win10	0.021	0.021	0.012	**5.081**
	(0.144)	(0.145)	(0.107)	
buyins won	1.460	1.479	0.902	**3.298**
	(29.752)	(30.219)	(8.427)	
mark-up[a]	17.149	13.231	17.164	-
	(10.760)	(4.622)	(10.775)	
percent requested[b]	55.748	46.923	55.783	-
	(16.705)	(14.367)	(16.706)	
percent staked[c]	52.766	0	52.984	-
	(19.749)	(0)	(19.497)	
winnings	96.743	96.082	115.980	-0.987
	(1054.002)	(1050.573)	(1149.310)	
lowbuyin	0.437	0.441	0.295	**18.225**
	(0.496)	(0.497)	(0.456)	
midbuyin	0.378	0.381	0.280	**12.760**
	(0.485)	(0.486)	(0.449)	
highbuyin	0.185	0.177	0.425	**-28.680**
	(0.389)	(0.382)	(0.494)	
weekend	0.488	0.481	0.690	**-25.618**
	(0.500)	(0.500)	(0.463)	
Observations	100,641	97,297	3,344	

Standard deviations appear in parentheses below the mean.

The t-statistics are from the null hypothesis that there is no difference between the unstaked mean and the staked mean, and that the variances of the two samples are unequal. T-statistics in **bold** are significant at the 5% level.

[a] *mark-up* has n = (3,351; 13; 3,338) for the (full; unstaked; staked) samples
[b] *percent requested* has n = (3,306; 13; 3,293) for the (full; unstaked; staked) samples
[c] *percent sold* has n = (3,169; 13; 3,156) for the (full; unstaked; staked) samples

11 Empirical Strategy

Ex ante, the impact of staking on performance is ambiguous. We propose two mechanisms through which staking could influence performance: alternative incentives created by staking, and a loosening of financial constraints faced by players. We discuss the theoretical predictions of each and describe the empirical tests we use to discern which impact dominates.

The practice of staking alters the incentives a player faces. As seen in Figure 10, a players marginal return to improving their finish position by one rank is lower when staked, as some percentage of the prize is reserved for the investor. Assuming that concentration/effort provision is costly for the player, muted incentives created by staking should lead the player to rationally choose a lower effort level when staked. Indeed, Ehrenberg and Bognanno (1990a,b) find that effort provision by professional golfers is lower when tournament prizes are smaller. Empirically, the impact of muted incentives should reveal itself in worse performance when staked. Further, if the marginal cost of effort is increasing, meaning that it is harder to maintain concentration the longer a tournament goes on, this impact should be increasingly evident in the right-tail of the performance distribution.

An argument for improved performance comes from the impact of staking on a players financial constraints. Given a finite budget and a choice among risky strategies, there can be a wedge between the optimal strategy for maximizing expected value in the current tournament and optimal strategy for maximizing the discounted expected value of all future tournaments played. The risk of ruin (a players chance of going broke) can dictate a lower variance, lower return strategy in the current tournament, than what is optimal given an unlimited budget. Staking increases a players effective budget, shrinking the gap between optimal strategies. This should allow a player to adopt a higher variance,

higher return strategy when staked. Empirically, this would reveal itself in terms of a higher return on investment when staked.

To summarize, staking could cause a change in performance due to (i) muted incentives leading to worse performance; or (ii) the relaxation of financial constraints leading to better performance. The first step in determining the dominant mechanism is to estimate the overall impact of staking on poker tournament performance. To do so, we employ a player by buyin-tier fixed effects model. This allows the comparison of a player's tournament outcomes when they are staked to their tournament outcomes when they are not staked. Additionally, by interacting the player fixed effects with three different tournament skill levels (proxied by buyin tiers), we allow for differences in a player's average performance/outcomes based on the skill level of the tournament.

Given that the payout structure of the tournaments is nonlinear, we explore several outcomes. The binary outcomes considered are *prize, smallwin, win3* and *win10. Prize* is an indicator equal to 1 if the player earns any prize, *smallwin* is an indicator for earning a prize of less than 3 times the tournament buyin, while *win3* and *win10* are indicators for a prize of greater than 3 and greater than 10 times the buyin, respectively. In addition to the binary outcomes we also look at the continuous outcomes *buyins won* (defined as the amount of the prize normalized by the entry fee) and *finishpercentile.*

The most straightforward prediction comes from examining *buyins won*, our measure of return on investment. If muted incentives reduce the effort level of players, we expect them to have a lower return on investment, whereas if staking allows a player to adopt a higher return strategy, this should appear in a larger return on investment. Similarly, players exerting less effort are less likely to attain the largest wins (*win3* and *win10*), especially if the marginal cost of effort is increasing in tournament duration, as these prizes are only awarded to those lasting to roughly the final 5% and 1% of all participants.

11.1 Monetary Outcomes

The first set of outcomes that we investigate are estimated with the following equation:

$$\mathbf{outcome}_{ibt} = \beta staked_{ibt} + \mu_{ib} + \mathbf{X}_t\mathbf{B} + \varepsilon_{ibt} \tag{3}$$

where $\mathbf{outcome}_{ibt} \in \{buyins\ won,\ win10,\ win3,\ smallwin,\ prize\}$ for player i, playing at buyin tier b, in tournament t. The variable $staked_{ibt}$ is a binary variable equal to 1 if a player was staked in a tournament and 0 otherwise, μ_{ib} is a vector of player by buyin-tier fixed effects, $\mathbf{X}$ is a vector of tournament specific control variables, and ε_{ibt} is a random error term.

The coefficient of interest is β, the coefficient on *staked*, which is the effect of being staked on the outcome observed. Under this fixed effects specification, β is consistently estimated so long as the only unobserved factors correlated with both *staked* and **outcome** do not vary over time. We estimate Equation (3) using ordinary least squares.[50]

11.2 Finish Position

In addition to the aforementioned monetary tournament outcomes, we also look at the rank that a player finishes in the tournament. A player that wins the tournament has a rank of 1, a player that finishes second has a rank of 2, and so on. The variable *finishpercentile* measures the percentile at which a player finishes the tournament:

$$finishpercentile = \left(1 - \frac{rank}{entries}\right) \times 100$$

[50]For the binary dependent variables (*prize, smallwin, win3, and win10*) we also estimate Equation (3) using a probit model and results are qualitatively identical.

and thus a higher *finishpercentile* is a better outcome for a player.[51]

The empirical model used to estimate staking's effect on a player's rank in a tournament is:

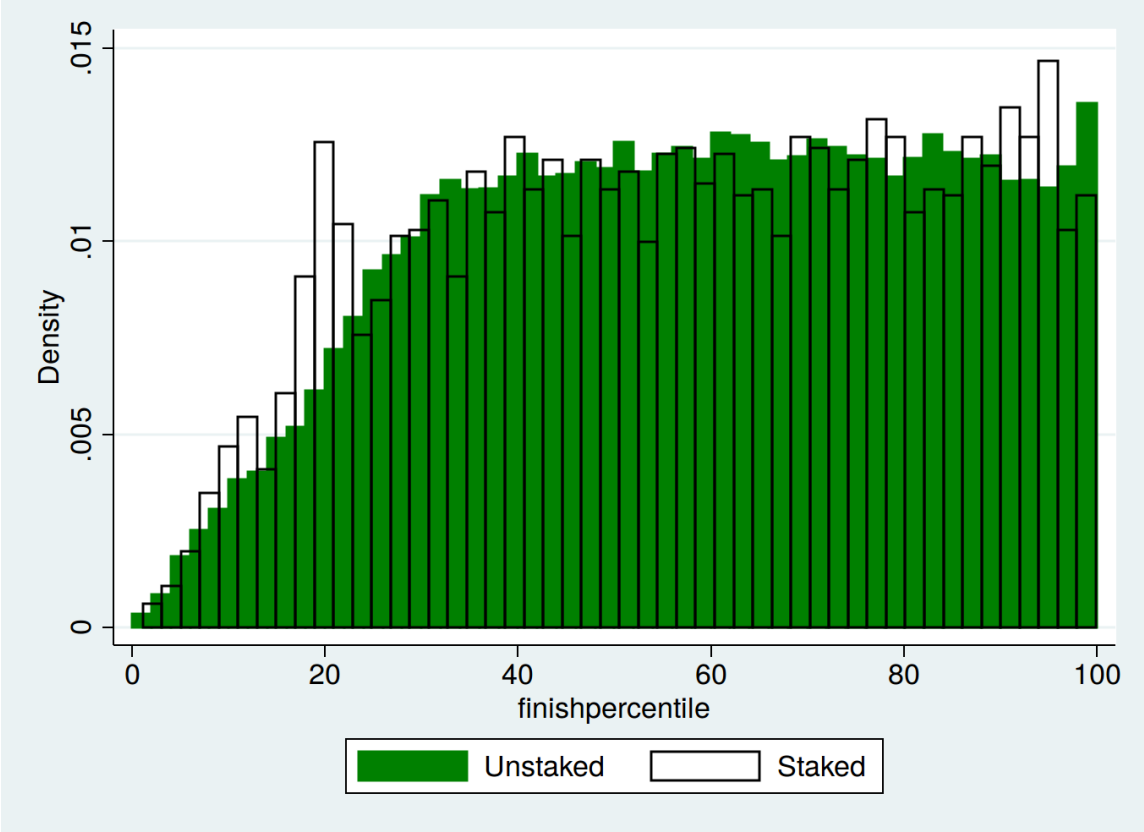
$$finishpercentile_{ibt} = \beta staked_{ibt} + \mu_{ib} + \mathbf{X}_t\mathbf{B} + \varepsilon_{ibt} \tag{4}$$

where $finishpercentile_{ibt}$ is the percentile of player i's finish at buyin tier b in tournament t, and $staked_{ibt}$, μ_{ib}, $\mathbf{X}_t\mathbf{B}$, and ε_{ibt} are as described in the previous subsection. Although β is consistently estimated by OLS when there are no unobserved variables correlated with both *finishpercentile* and *staked*, we also consider quantile regression (QREG) as an estimation method. As displayed in the histogram of *finishpercentile,* Figure 12, it appears that the effect of staking on finishing position is concentrated in the right-tail of the distribution of *finishpercentile.* Therefore, our preferred method of estimation of Equation (4) is QREG since it allows for changes in an expected quantile rather just the expected mean.[52]

[51]Note that, as constructed, *finishpercentile* is biased downward - it never takes on a value of 100; an alternative measure is also considered where *finishpercentile* is biased upward - it can never be zero.

[52]We also perform OLS on Equation (3) and present it in Table 12 as a baseline.

Figure 12: Distribution of *finishpercentile*: Staked versus Unstaked



12 Results

12.1 Main Results

12.1.1 Monetary Outcomes

Table 11 shows the ordinary least squares (OLS) results for estimating the effect of being staked on the set of monetary outcomes. We compare a player's outcomes at a specific buyin-tier when they are staked to when they are not staked. We find that staking has a negative effect on a player's average normalized winnings from a poker tournament. Our results also suggest that the probability of a player winning a large

prize are lower when playing in a tournament in which they are staked. While the probability of a large prize is decreasing, staking appears to increase the probability of a small prize. Our preferred estimates also suggest that the overall probability of returning any prize is not affected by a player receiving staking.

We first focus on a continuous measure of poker tournament prizes and its relationship to whether or not a player has been staked for a poker tournament. The dependent variable in Column 1 of Table 11, *buyins won*, scales the prize won in a tournament by the entry fee for that tournament. The results here suggest that being staked decreases, on average, the number of *buyins won* by 0.67. This is a 44.42 percent decrease when compared to the number of *buyins won* by the same player, playing in the same tournament-skill-tier, when unstaked. This result provides evidence that staking, by muting a player's incentives, has a negative effect on the monetary outcomes of poker players.

We now turn our attention to columns 2 and 3. The results found in these columns provide more evidence that staking mutes the incentives of the players in our sample. The dependent variables in columns 2 and 3, *win10* and *win*3, are two different binary measures of large poker tournament wins. The results in columns 2 and 3 suggest that when a player receives staking for a tournament, that player's probability of winning at least ten times the amount of the entry fee decreases by 0.68 percentage points, while their probability of winning at least three times the amount of the entry fee decreases by 0.96 percentage points. These are 32.33 percent and 19.49 percent reductions, respectively. Thus, when a player is staked, they are less likely to win a large poker tournament prize.

Table 11: Monetary Outcomes

VARIABLES	(1) *buyins won*	(2) *win10*	(3) *win3*	(4) *smallwin*	(5) *prize*
staked	-0.665***	-0.00679***	-0.00955**	0.0137**	0.00465
	(0.183)	(0.00225)	(0.00391)	(0.00584)	(0.00680)
Tournament Characteristics:					
buyin	-0.000939*	-6.16e-06	-7.90e-09	8.44e-06	8.50e-06
	(0.000486)	(5.34e-06)	(9.36e-06)	(1.25e-05)	(1.51e-05)
entries	-0.000843	-3.24e-06***	-1.46e-06**	3.64e-06***	2.03e-06**
	(0.00102)	(3.54e-07)	(6.42e-07)	(7.34e-07)	(9.29e-07)
$entries^2$	4.66e-08	8.30e-11***	7.65e-11***	-7.57e-11***	0
	(5.00e-08)	(0)	(0)	(0)	(0)
unlimited rebuy	0.645***	0.00937***	0.0186***	0.0178***	0.0369***
	(0.170)	(0.00159)	(0.00231)	(0.00279)	(0.00347)
single rebuy	-0.113	0.00685**	0.00827**	0.0155***	0.0240***
	(0.428)	(0.00277)	(0.00382)	(0.00489)	(0.00598)
fast	0.0423	-0.00647***	-0.0115***	0.00135	-0.0105***
	(0.417)	(0.00119)	(0.00180)	(0.00247)	(0.00296)
slow	0.290	-0.00599**	-0.00912**	-0.00961	-0.0178**
	(0.620)	(0.00265)	(0.00454)	(0.00624)	(0.00748)
Observations	100,641	100,641	100,641	100,641	100,641
R-squared	0.012	0.007	0.007	0.006	0.008

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All regressions include player by buyin-tier fixed effects and indicators for whether or not the tournament was part of a special tournament series, or played on the weekend.

Focusing on Column 4 from Table 11, we find that being staked increases the probability that a player wins a prize of less than three times the entry fee by 1.37 percentage points - a 16.12 percent increase when compared to the average probability of a "small win". The result found in Column 5, where *prize* is the dependent variable, finds an imprecisely estimated positive coefficient on *staked*. The 95 percent confidence interval on the estimate of β ranges from -0.0087 and 0.0180. Thus, the effect of *staked* on *prize*

ranges from a 6.52 percent decrease to an 13.5 percent increase. Looking over this range of possibilities, we conclude that staking has a small, if not zero, effect on whether or not a player wins a prize.

Combining all of the results in Table 11, our conclusion is that staking has a negative effect on a poker player's tournament outcomes. Although staking possibly has a very small positive effect on whether or not a player wins any prize, we see that staking is negatively related with *buyins won*, *win10*, and *win3*. Further, when we add the evidence that smaller poker tournament wins are more likely when a player is staked, it appears that staking has the effect of reallocating large poker tournament wins to smaller poker tournament wins. When we combine this explanation with the negative coefficient on staked in the *buyins won* specification (implying a negative change in the average tournament prize), it appears that staking has an adverse effect on the distribution of poker tournament outcomes. To expand upon this idea, we turn our attention to a comparison of the rank at which a player finishes a tournament under the two staking scenarios.

12.1.2 Finishing Rank Outcomes

In addition to monetary outcomes, we also examine staking's effect on the rank at which a player finishes in a poker tournament. Since poker tournaments vary in size even within a given buyin level, as described above, we create a measure of the percentile at which a player finishes a tournament. Initial evidence that staking has an effect on where a player finishes in a tournament is found in Figure 12. This figure shows staked players are less likely to finish in the extreme upper percentiles. For a reference point, and tying this figure to the monetary outcomes above, the extreme upper percentiles are where the large poker tournament prizes are won.

Table 12 shows the results of regressing *finishpercentile* on an indicator for whether or not a player received staking for that poker tournament. In Column 1 we use ordinary least squares (OLS) to estimate this relationship and find a positive relationship between whether or not a player is staked and the percentile at which a player finishes a tournament. Although the OLS results suggest that staking has almost no impact on a player's poker tournament ranking outcome, there is reason to believe that comparing the average *finishpercentile* of a player when they are staked to when they are not staked is not informative about the effect that staking has on a player's incentives. Columns 2 through 7 are the results of regressing *finishpercentile* on *staked* using quantile regression (QREG) as the estimation method. The negative effect that staking has on a player's incentives is found in the right-tail of the distribution of *finishpercentile* - where there are poker tournament prizes. On average, poker tournaments begin paying out players around the 90th percentile. The QREG results show that, from the 90th quantile on, staking has a uniformly negative effect on the expected *finishpercentile* of a player. Again, this suggests that the negative effect on poker tournament outcomes generated by muted incentives dominates any positive effect on these outcomes generated by loosening a player's financial constraints.

Table 12: Finishing Position Outcomes

	Dependent Variable: *finishpercentile*						
Estimation	OLS	Quantile					
Quantile	N/A	25	50	75	90	95	97.5
staked	0.0565	-0.345	-0.389	0.413	-0.0595	-0.726*	-0.450*
	(0.499)	(0.693)	(0.768)	(0.731)	(0.556)	(0.385)	(0.246)
Tournament Characteristics:							
truebuyin	-0.00384***	-0.00896***	-0.00425**	-7.34e-05	-0.000318	0.000629	-0.000732
	(0.00122)	(0.00157)	(0.00174)	(0.00165)	(0.00126)	(0.000871)	(0.000557)
entries	-3.04e-05	-0.000199**	-7.91e-05	6.24e-05	-5.66e-05	-4.36e-05	-4.79e-05
	(6.66e-05)	(9.09e-05)	(0.000101)	(9.59e-05)	(7.30e-05)	(5.05e-05)	(3.23e-05)
$entries^2$	-7.07e-10	9.90e-10	0	7.60e-10	2.79e-09	1.91e-09	1.18e-09
	(1.88e-09)	(2.42e-09)	(2.68e-09)	(2.55e-09)	(1.94e-09)	(1.34e-09)	(8.59e-10)
unlimited rebuy	7.756***	11.66***	8.929***	5.318***	2.201***	1.107***	0.540***
	(0.220)	(0.332)	(0.368)	(0.350)	(0.266)	(0.184)	(0.118)
single rebuy	5.158***	8.376***	5.977***	3.810***	1.787***	0.825**	0.412*
	(0.432)	(0.606)	(0.672)	(0.639)	(0.486)	(0.336)	(0.215)
fast	-1.650***	-1.033***	-5.298***	-2.504***	-1.074***	-0.674***	-0.455***
	(0.217)	(0.314)	(0.348)	(0.331)	(0.252)	(0.174)	(0.111)
slow	0.222	-0.377	0.122	-0.235	-0.655	-0.663	-0.0764
	(0.552)	(0.810)	(0.898)	(0.854)	(0.650)	(0.450)	(0.288)
Observations	100,641	100,641	100,641	100,641	100,641	100,641	100,641
R-squared	0.037						

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All regressions include player by buyin-tier fixed effects and indicators for whether or not the tournament was part of a special tournament series, or played on the weekend.

Both Figure 12 and the QREG results provide evidence that when a player receives staking there is a negative relationship with where that player finishes in a poker tournament. This impact is concentrated in the right-tail of the distribution. One reason that we may observe this, is that the marginal cost of effort is increasing in the duration of a tournament, as it becomes progressively challenging/costly to maintain concentration.

12.2 Addressing Issues of Selection Bias

The results described in the previous subsection imply that the negative effect of staking, the reduced incentives for marginal effort, tend to dominate the positive effect of staking, the loosening of a player's financial constraints. One concern is that there is a sample selection issue that would yield the same results. For example, a player who has private information that their recent results appear better than their true skill level (lots of recent luck) may seek staking with a profit motive. By seeking staking at markup, they could ensure a guaranteed payout that is perhaps higher than their true expected value in a given tournament. Our initial results measure the total impact of staking on performance, this includes the direct impact of staking, and any within-player selection effects. To address the possibility that selection is biasing our results towards worse performance when staked, we employ three tests.

We first examine visual evidence regarding a player's poker tournament outcomes before they sought out staking and compare this to tournament outcomes directly after this first incident of staking. Here, profit motive should reveal itself in unusually strong performance directly before staking. We then look at how variation in the amount that a player was staked effects a player's poker tournament outcomes only for tournaments where a player is staked. This allows us to allay concern about bias caused by sample selection as the comparison group is limited to other staked tournaments of the same

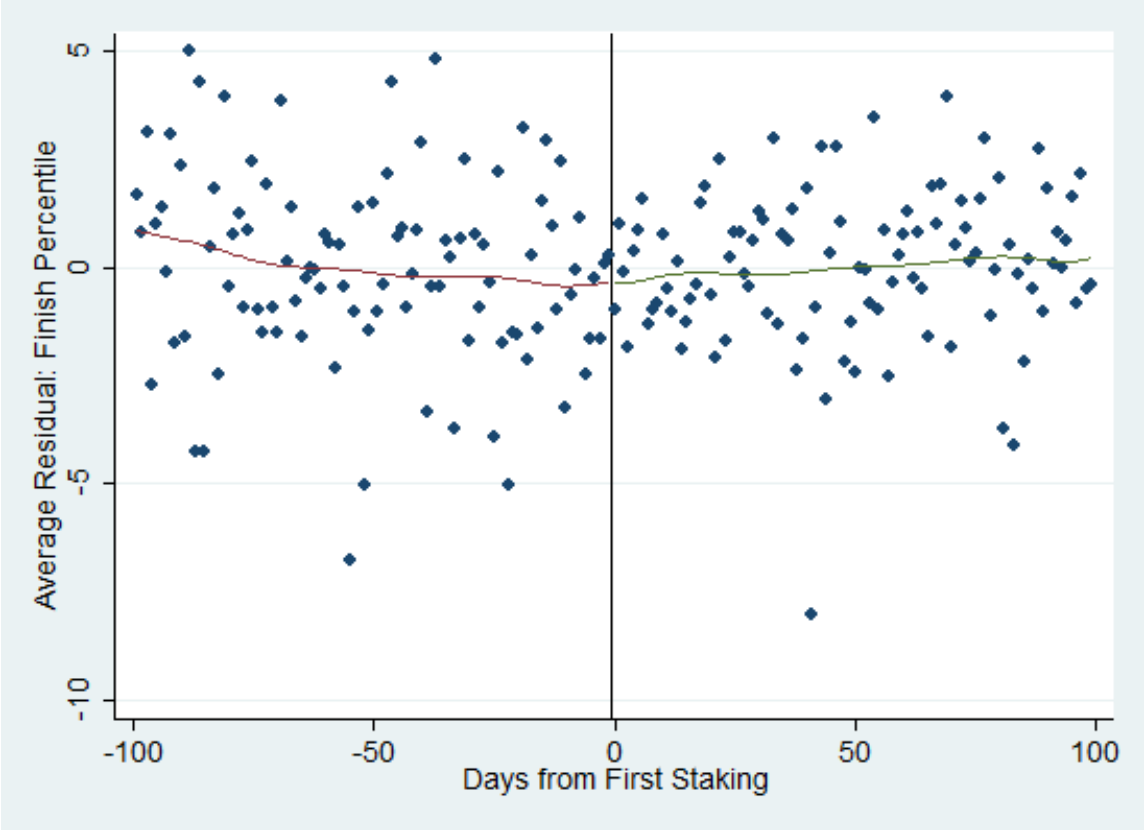
buyin tier by the same player. Finally, we consider only variation in percent staked generated by take-up on the investor side (players not receiving their full staking request).

12.2.1 Staking Timeline

One scenario that would diminish our finding, that staking has a negative impact on relatively large poker tournament prizes, would be if poker players had abnormal positive results right before they sought out staking and then returned to "normal" immediately, and coincidentally, right after receiving staking. To investigate this potential scenario we regress *finishpercentile* on month dummies and all of the controls found in Equation (3) for all of a player's unstaked tournaments. We then average the residuals by each day pre- and post- a player's first day of staking. Plotting these residuals in Figure 13, we see that there is no long positive run of residuals prior to a player's first day of staking. This lack of a pattern in the residuals mitigates the concern that our findings are due to a player returning to their average poker tournament outcomes after a lucky streak.[53]

[53] We also perform a Chow test (untabulated) for a structural break at the time of a player's first staked tournament and include a cubic time trend. The results of the test fail to reject the null hypothesis of no structural break.

Figure 13: Residual Plot of *finishpercentile* Around the First Staking Incident



Notes: The residuals are generated from a regression of finishpercentile on all standard controls and month fixed effects, using a sample of only unstaked tournaments. Each point is the average of all residuals for that day relative to the first instance of staking. Fitted lines are results of locally weighted regression.

12.2.2 Percent Staked

The empirical model used to test how *percent staked* affects tournament outcomes is:

$$\mathbf{outcomes}_{ibt} = \alpha \textit{percent staked}_{ibt} + \mu_{ib} + \mathbf{X}_t\mathbf{B} + \varepsilon_{ibt} \tag{5}$$

where $\mathbf{outcomes}_{ibt} \in \{\textit{buyins won, win3, win10, smallwin, prize, finishpercentile}\}$, $\textit{percent staked}_{ibt}$ is the amount of staking that player i playing at buyin-tier b received

for tournament t, and the remainder of the variables are as defined above. We use OLS to estimate *buyins won* and the binary dependent variables, whereas we use QREG to estimate *finishpercentile*.

The coefficient of interest is α, the coefficient on the variable *percent staked*. Since Equation (5) is also a player by buyin-tier fixed effects specification, α is consistent so long as there are no time variant unobserved variables correlated with the outcome in question and the *percent staked*. As the amount of staking that a player receives increases we expect that their incentives become more and more muted, *ceteris paribus*. Hence, a positive $\hat{\alpha}$ is consistent with the incentive mechanism story when the dependent variables are *smallwin*, and a negative $\hat{\alpha}$ for *buyins won, win10*, *win3*, and *finishpercentile*.

Table 13 shows how the amount of staking received for a tournament impacts the aforementioned monetary outcomes. We restrict our sample to only staked tournaments where the precise amount of staking is known.[54] As the amount of staking a player receives (*percent staked*) increases, we find the average number of buyins won decreases and that the likelihood of a relatively large win (as measured by *win10* and *win3*) also decreases. When we look at the probability of winning any prize (*prize*) or winning a relatively small prize (*smallwin*), our results suggest that there is essentially no effect of staking on these outcomes. The results in Table 13 provide evidence that our main results are not being driven by sample selection, but by changes in a player's incentives. This evidence is strengthened by the choice of our sample - we only consider observations where a player has staking for a poker tournament. We find results consistent with the premise that when the percentage of staking that a player receives increases, the marginal return for a player's effort decreases and the player chooses a lower effort level accordingly. The poker tournament outcome for a player, given these muted incentives,

[54] We also investigate the full sample of poker tournaments by using the sample of players with known staking amounts and setting the unstaked players *percent staked* equal to zero. The results are similar to Table 13, except with coefficients that are smaller in magnitude (except for the coefficient on *percent staked* when the dependent variable is *smallwin*) but more precisely estimated.

is a decreased likelihood of winning large poker tournament prizes.

Table 13: Monetary Outcomes - Percent Staked (Restricted Sample)

VARIABLES	(1) *buyins won*	(2) *win10*	(3) *win3*	(4) *smallwin*	(5) *prize*
percent staked	-0.0371*	-0.000352**	-0.000620*	0.000165	-0.000435
	(0.0225)	(0.000172)	(0.000359)	(0.000590)	(0.000668)
Tournament Characteristics:					
buyin	0.000234	7.86e-06	3.44e-05	2.84e-05	6.24e-05
	(0.000438)	(1.51e-05)	(2.84e-05)	(3.46e-05)	(4.44e-05)
entries	-5.70e-05	-3.72e-06	-1.85e-07	7.27e-06*	7.17e-06
	(5.66e-05)	(2.50e-06)	(3.64e-06)	(4.10e-06)	(5.21e-06)
$entries^2$	3.03e-09	2.33e-10	1.71e-10	-2.11e-10**	-0
	(2.05e-09)	(1.58e-10)	(1.71e-10)	(1.03e-10)	(1.88e-10)
unlimited rebuy	-0.114	-0.00397	0.0161	-0.00121	0.0152
	(0.347)	(0.00611)	(0.0129)	(0.0175)	(0.0210)
single rebuy	-0.363	-0.00395	0.0168	0.0627**	0.0795**
	(0.351)	(0.00901)	(0.0191)	(0.0310)	(0.0350)
fast	-1.188**	-0.0134**	-0.0314**	-0.0336	-0.0644**
	(0.502)	(0.00523)	(0.0141)	(0.0267)	(0.0294)
slow	0.307	0.0132	0.0116	0.0141	0.0265
	(0.434)	(0.0134)	(0.0218)	(0.0334)	(0.0379)
Observations	3,169	3,169	3,169	3,169	3,169
R-squared	0.028	0.043	0.065	0.060	0.066

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All regressions include player by buyin-tier fixed effects and indicators for whether or not the tournament was part of a special tournament series, or played on the weekend

We also examine the effect that *percent staked* has on the rank at which a player finishes in a poker tournament. The results, which are found in Table 14, suggest that as the amount of staking a player receives increases, that player has a worse expected poker tournament ranking. As with the main results, this is made evident when looking at the

extreme right-tail of the distribution of *finishpercentile*.[55] These results are consistent with the idea that the alternative payout structure created by staking mutes player incentives and leads to worse poker tournament outcomes.

[55] As with the monetary outcomes, we also expand our sample to include all players whose staking amount is known by setting *percent staked* equal to zero for all unstaked players. The results are qualitatively similar to the restricted sample results, however none of the coefficients are precisely estimated.

Table 14: Finishing Rank Outcomes - Percent Staked (Restricted Sample)

	Dependent Variable: *finishpercentile*						
Estimation	OLS	Quantile Regression					
Quantile	N/A	25	50	75	90	95	97.5
percent staked	0.00799	-0.00351	0.0368	0.00263	-0.0547	-0.0192	-0.0301*
	(0.0488)	(0.0650)	(0.0693)	(0.0651)	(0.0415)	(0.0283)	(0.0154)
Tournament Characteristics:							
buyin	0.00658*	0.00453	0.0107**	0.00792*	0.00394	0.00442**	0.00310***
	(0.00340)	(0.00480)	(0.00512)	(0.00481)	(0.00306)	(0.00209)	(0.00113)
entries	0.000520	0.000624	0.000546	0.000553	0.000188	-5.93e-05	-2.22e-05
	(0.000381)	(0.000525)	(0.000560)	(0.000526)	(0.000335)	(0.000229)	(0.000124)
$entries^2$	-1.47e-08	-2.36e-08	-9.71e-09	-8.37e-09	-1.52e-09	3.60e-09	2.81e-09
	(1.41e-08)	(1.79e-08)	(1.91e-08)	(1.79e-08)	(1.14e-08)	(7.79e-09)	(4.23e-09)
unlimited rebuy	6.509***	10.59***	8.155***	3.949*	0.509	0.227	-0.123
	(1.473)	(2.260)	(2.410)	(2.264)	(1.441)	(0.984)	(0.534)
single rebuy	6.400***	9.784***	3.899	5.979*	5.692***	2.486*	2.709***
	(2.371)	(3.296)	(3.515)	(3.301)	(2.102)	(1.435)	(0.779)
fast	-6.368***	-2.549	-9.600**	-5.695	-3.147	-2.505	-2.521***
	(2.395)	(3.582)	(3.820)	(3.588)	(2.285)	(1.560)	(0.847)
slow	4.064	5.748	5.590	3.243	-0.259	-1.542	-1.861**
	(2.687)	(3.996)	(4.261)	(4.002)	(2.549)	(1.740)	(0.945)
Observations	3,169	3,169	3,169	3,169	3,169	3,169	3,169
R-squared	0.098						

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All regressions include player by buyin-tier fixed effects and indicators for whether or not the tournament was part of a special tournament series, or played on the weekend.

12.2.3 Staking Gap

Another way that we measure how staking changes the incentives of poker players is through the difference between how much staking a player received for a poker tournament and how much staking a player sought out for that same poker tournament. This allows us to look at variation generated by differences in how much a player was willing to sell of their winnings and how much they would actually have to give up. Thus, we create the variable *gap* for player i at buyin-tier b in tournament t, which we define as:

$$gap_{ibt} = percent\ requested_{ibt} - percent\ staked_{ibt}$$

As this difference increases, a player's marginal return to finishing position also increases, compared to what they were willing to give up to their investors, *ceteris paribus*. Therefore, we expect the average number of *buyins won* to increase and the probability of large poker tournament wins to increase as *gap* increases.[56]

To investigate the relationship between the monetary outcomes and *gap* we introduce the following specification:

$$\mathbf{outcome}_{ibt} = \lambda_1 gap_{ibt} + \lambda_2 percent\ requested_{ibt} + \mu_{ib} + \mathbf{X}_t\mathbf{B} + \varepsilon_{ibt} \tag{6}$$

where $\mathbf{outcome}_{ibt} \in \{buyins\ won,\ win10,\ win3,\ smallwin,\ prize\}$, gap_{ibt} is as defined above, $percent\ requested_{ibt}$ is the percentage of winnings that player i offered to potential buyers for tournament t at buyin-tier b, and μ_{ib}, $\mathbf{X}_t\mathbf{B}$, and ε_{ibt} are as described above. This specification allows us to compare players with different levels of *gap*, but who requested the same amount of staking. Since we include the amount of staking that

[56] We consider two versions of *gap*, one where *gap* is restricted to non-negative numbers and one where *gap* does not have this restriction. For *gap* to be negative, a player must receive more staking than initially requested. We find that both definitions yield qualitatively similar results.

the player requested when they posted their advertisement on twoplustwo.com, concerns regarding adverse selection are further mitigated. Thus, the estimate of λ_1, the coefficient on *gap*, will provide more insight into how incentives play a role in the effort and concentration choices of poker players.

Estimating Equation (6) by OLS yields the results found in Table 15. The sample is restricted to only players that received staking for a poker tournament and whose staking gap was known.[57] Columns 1 through 3 provide further evidence that as a player's incentives are increased, as measured by an increase in *gap*, their average number of buyins won increases, as does the probability of player achieving a large win (*win10* and *win3*), *ceteris paribus*.[58] The results in Columns 4 and 5 are consistent with our other findings regarding smaller poker tournament wins (*smallwin*) and the probability of winning any prize (*prize*). As *gap* increases, the probability of a small win decreases and the probability of winning any prize remains unchanged. With the exception of the specification where *win10* is the outcome variable, the coefficients on the variable *gap* are imprecisely estimated. However, when considering the outcomes of *buyins won*, *win10*, and *win3*, this lack of statistical significance is most likely due to a negative bias from (potential) investors being able to observe something about the poker player that the econometrician cannot. Specifically, we propose that the aforementioned unobserved variable would be positively correlated with *gap* and negatively correlated with *buyins won* or either of the big win dummy variables.[59] Since, our estimated coefficients on *gap* for these specifications are positive, while the potential bias is negative, this implies that the incentives are the dominant factor with respect to the effects of staking on poker

[57] We lose 26 observations due to the advertisement not listing how much staking a player was seeking.

[58] Additionally, the estimated coefficients in Table 15 are similar in absolute value to the estimates when the monetary outcomes are regressed on only *percent staked* and the controls (i.e. the results found in Table 13).

[59] For example, suppose that a potential investor thinks that a poker player has gotten lucky in their successful poker tournaments. This investor is less likely to invest in said player, and if the investor is correct regarding the players tournaments outcomes being lucky results, this player will eventually revert to their mean poker tournament prizes in future tournaments.

tournament performance.

Table 15: Monetary Outcomes (Staking Gap and Percent Requested; Restricted Sample)

VARIABLES	(1) *buyins won*	(2) *win10*	(3) *win3*	(4) *smallwin*	(5) *prize*
gap	0.0304	0.000285*	0.000221	-0.000113	0.000100
	(0.0203)	(0.000166)	(0.000437)	(0.000704)	(0.000805)
percent requested	-0.0479	-0.000458	-0.00117**	0.000152	-0.000983
	(0.0395)	(0.000295)	(0.000530)	(0.000868)	(0.000978)
Tournament Characteristics:					
buyin	0.000338	8.88e-06	3.98e-05	2.87e-05	6.80e-05
	(0.000586)	(1.58e-05)	(2.93e-05)	(3.54e-05)	(4.55e-05)
entries	-5.78e-05	-3.75e-06	-1.91e-07	7.41e-06*	7.30e-06
	(5.71e-05)	(2.51e-06)	(3.67e-06)	(4.13e-06)	(5.25e-06)
entries^2	3.06e-09	2.33e-10	1.71e-10	-2.12e-10**	-0
	(2.05e-09)	(1.58e-10)	(1.71e-10)	(1.04e-10)	(1.88e-10)
unlimited rebuy	-0.119	-0.00404	0.0160	-0.000888	0.0156
	(0.352)	(0.00617)	(0.0130)	(0.0176)	(0.0212)
single rebuy	-0.373	-0.00404	0.0164	0.0630**	0.0794**
	(0.349)	(0.00906)	(0.0193)	(0.0313)	(0.0352)
fast	-1.193**	-0.0134**	-0.0315**	-0.0339	-0.0648**
	(0.502)	(0.00527)	(0.0142)	(0.0269)	(0.0296)
slow	0.308	0.0135	0.0114	0.0163	0.0287
	(0.442)	(0.0136)	(0.0223)	(0.0340)	(0.0386)
Observations	3,143	3,143	3,143	3,143	3,143
R-squared	0.028	0.043	0.065	0.060	0.066

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

All regressions include player by buyin-tier fixed effects and indicators for whether or not the tournament was part of a special tournament series, or played on the weekend

13 Conclusion

Performance pay, broadly defined, was used in 39% of jobs in the US private sector during 2013 (Gittleman and Pierce, 2013). Despite this prevalence, particularly among the highest quartile of wage-earners, empirical studies are relatively few. To assess the impact of alternative compensation on performance, we make use of a unique setting, the advent of a formal staking market for online poker players. This market allows poker players to move from a pure variable payment–based on tournament performance–to a hybrid structure with a fixed fee and variable tournament component. Our central finding is that the performance is significantly better under the pure variable payment scheme, with poker tournament earnings falling by 44% under the hybrid scheme.

To address the concern that selection bias could be contributing to our finding, we run three empirical tests. First, we examine unstaked tournament results surrounding a players first staking event and find no significant performance differences. Then, limiting the sample to only tournaments with nonzero staking and using variation from the intensive margin, percent staked, we find that performance improves as more weight is given to the variable component of pay. Finally, we examine variation in percent staked created on the investor side of the market. While these estimates are noisy, earnings are higher when a player receives less than their desired level of staking, *ceteris paribus*. The balance of evidence suggests that the worse performance found under staking is primarily due to lowering the marginal return to an increase in rank, rather than through selection.

Our results confirm the prediction of tournament theory that larger prizes induce higher effort levels from competitors. This has implications for firms, where promotions often follow a tournament structure with employees promoted based on their performance relative to other employees. Our results suggest that increasing the prize (the

value of the promotion) can be an effective way to increase productivity. Tournament theory also suggests that the higher the variance in the mapping between effort and output, the less impact tournament prizes will have on effort levels (Lazear and Rosen, 1981; Eriksson, 1999). Despite the high variance in poker tournament outcomes, we still find economically meaningful impacts from varying tournament prizes. This suggests that tournament incentives can still play an important role in industries where output is highly variant.

Chapter III
Scanning the Tables: Do Professional Online Poker Players Respond to Earning Conditions?

14 Introduction

Dynamic models of lifecycle labor supply predict that when facing a transitory wage shock, workers substitute intertemporally between labor and leisure, working more during high wage periods and reducing hours worked during low wage periods when the opportunity cost of leisure is relatively low (Lucas and Rapping, 1969). Despite this straightforward prediction, previous empirical studies have found little evidence of intertemporal substitution of labor. Estimates using individual panel data suggest that this elasticity is quite small, or perhaps negative (MaCurdy, 1981; Browning, Deaton, and Irish, 1985; Altonji, 1986).

One concern with this set of studies, is that they relate annual changes in hours worked to annual changes in average wages. It seems unlikely that annual wage changes are purely transitory, and hence they would be associated with nontrivial income effects. If there is any increase in expected lifetime wealth, this would bias estimates of the intertemporal substitution elasticity downwards, and could explain the very small empirical estimates. Another potential concern is that there could be constraints on workers' labor supply. Indeed, there is strong evidence that in many industries workers are not free to set their own working hours (Dickens and Lundberg, 1993; Farber, 2005). To address these concerns, a new strand of the literature has focused on unconventional jobs that conceivably have flexibility in hours worked and purely transitory wage shocks.

These studies have focused on taxi cab drivers (Camerer et al., 1997; Farber, 2005;

Crawford and Meng, 2011; Farber, 2014) who can choose when to quit a shift, and stadium vendors (Oettinger, 1999) and bike messengers (Fehr and Goette, 2007), who can choose the number of shifts worked. Although early results in the area were mixed, with estimated elasticities ranging from -1 to 1.5, a majority find that there is a significant positive intertemporal labor supply elasticity, as predicted by neoclassical theory. While these studies represent a significant step forward from their predecessors, each category of workers still faces limitations on their labor supply. Taxi drivers must pay a significant fixed cost to rent their medallion, can only choose when to quit while on shift, and are constrained to work a maximum of 12 hours. Stadium vendors and bike messengers choose whether to participate, but face fixed shift lengths and set game dates or shift times. Further, these studies focus on low-wage jobs in localized regions (individual cities), bringing up concerns about external validity. I overcome these challenges by analyzing a unique international labor market with relatively high wages and the ability to start and stop work at any time.

Matching hours worked data from the *Online Poker Database of the University of Hamburg* with a dataset I assembled on daily poker earnings, I analyze the labor supply decisions of professional online poker players. This is an attractive market for studying labor supply because players have complete flexibility in hours worked and can adjust effort on an observable margin, altering the number of tables played simultaneously.[60] Additionally, and crucial for estimation, expected wages for a professional player vary greatly across time.

When a pro sits down at a poker table, their expected wage is inversely related to the skill level of their opponents. At a table with many inexperienced recreational players, expected wages are high. At a table with only skilled professionals, expected wages are low and could even be negative. Online professionals often play at many

[60]The setting in Fehr and Goette (2007) allows for inference of effort, based on revenue earned during a fixed-length shift.

tables at once, and so their wage expectation depends on the overall composition of the player pool for the particular game type and stake they are playing. Ceteris paribus, an increase in the number of recreational players raises expected wages for pros, while a decline in the number of recreational players reduces expected wages. Using variation in expected wages created by the amateur portion of the player pool, I find that the top professionals respond by playing more often and adding additional tables when expected wages are high. In contrast, weaker professionals only respond by adding additional tables, suggesting that the skills that make a poker player successful are related to optimizing work decisions.

The remainder of this proposal is organized as follows. Section 2 provides the background about online poker cash games that is necessary for understanding the empirical analysis. Section 3 discusses the two data sources, detailing how they are merged and what unique variables are provided in each. Section 4 outlines the empirical framework and central estimation equations, while Section 5 presents the central results, including tests for heterogeneity across player type. Section 6 concludes with a discussion of the implications of the findings.

15 Online Poker Cash Games

To participate in an online poker cash game, a prospective player must first create an account associated with the hosting website and deposit money for use at the tables. During the sample period, the only restriction on account creation was that the individual must be of legal gambling age in their jurisdiction. Upon account creation, the prospective player chooses a unique username that identifies them whenever playing. To participate in a particular cash game, the player must choose a table from the playing lobby with an available seat (the virtual poker tables used in this analysis are limited to

6 players) and select "take-seat", to join the game.

When joining the game, money is transferred from the players virtual account with the hosting site to the table. A player may choose the amount of money to put on the table, subject to restrictions based on the betting limits of the game. For the game studied here, the player may start with anywhere from 20 times the minimum bet to 100 times the minimum bet.[61] After taking a seat, a player is a part of the game and will be dealt into the next hand. Play proceeds according to the rules of the game, and any chips won or lost are added to the players virtual chipstack at the table. All chips are denoted at their dollar value, and importantly for the premise of this paper, a player may get up after any hand and leave with the amount of money they possess.

The specific poker game analyzed in this paper is No-Limit Texas Hold'Em, by far the most prevalent type of poker played during the study period.[62] Levitt, Miles, and Rosenfield (2012) and Levitt and Miles (2014) provide convincing empirical evidence documenting the importance of skill in Texas Hold'Em. Notably, when participating in World Series of Poker tournaments, *a priori* identified pros realized a return on investment that was 46 percentage points higher than nonpros. In my analysis, I analyze the labor supply decisions of professionals as defined based on high playing volume and success over a nine month pre-period.

16 Data Description

The primary data source for this analysis is the *Online Poker Database of the University of Hamburg* (OPD-UHH) which contains high-frequency labor supply data for

[61] In between hands a player may always add additional funds to bring their total allotment of chips up to 100 times the minimum bet. However, a player may not remove chips from the table and remain seated. The only way to remove chips from the table is to leave that game.

[62] No Limit Texas Hold'Em comprises over 85% of all hours played during my sample period.

the universe of players on one major online poker site - Full Tilt Poker.[63] To create this dataset, the OPD-UHH took a virtual "snapshot" of the lobby for Full Tilt Poker at approximately 10-minute intervals. Each snapshot provides a list of players playing at each table (identified by username), the player's country of origin, what the stakes of the table are, and what the game type is. The data spans six months, from Sept 6th, 2009 to March 11th, 2010. However, I use the September period solely to estimate the skill level of all players in the sample. Further, due to the advent of new game offerings in 2010 that were untracked, I limit my labor supply study period to Oct 1st - Dec 18th 2009.[64] Due to the substantial time costs associated with cleaning the labor supply data, I focus on one particular game type and stake: \$1/\$2 No-Limit Texas Hold'Em 6-Max.[65]

In order to define professional players and to create a high-frequency measure of earning conditions, I augment the labor supply data with daily earnings data procured from a reliable poker datamining company. The datamining company tracked the vast majority of hands occurring on Full Tilt from Jan 2009-Dec 2010.[66] This dataset tracks earnings for each player at each unique game type by stake combination, at the daily level. For example, Player A would have one observation for Jan 14th 2009, playing \$1/\$2 (stake)[67] No Limit-Holdem (game type) and another observation on Jan 14th

[63]The OPD-UHH tracked this information across multiple poker sites, but with about 20% of the market share during the study period, FullTilt is by far the largest site for which earnings data is also available.

[64]I stop the sample on Dec. 18th rather than Dec. 31st, because playing volume is significantly different around the winter holidays, and unobserved differences in religious observance could alter the value of leisure for different players at different times. I also omit the Wednesday - Sunday surrounding Thanksgiving from my analysis, as a plurality of professionals are U.S. based.

[65]This was the most popular cash game played during the study time frame with a full buyin of at least \$100. Further, this mid-stake level is preferred to higher stakes games where scarcity of tables could constrain a players choice of when or how many tables to play. Finally, recreational players make up a larger portion of the player pool at lower stakes and this analysis uses plausibly exogenous variation in the recreational portion of the player pool.

[66]Coverage is 95%+ and the missing data generally occurs when FullTilt updated their software and lasts until a fix was made by the datamining company.

[67]Stakes refer to the betting limits of the game. The second number, \$2, refers to the big blind. This is the smallest increment that can be bet in a single action. For reference, a standard buy in is 100 times the big blind, or \$200 for a \$1/\$2 game.

2009, playing \$2/\$4 (stake) Pot Limit Omaha (game type).

Because the earnings data is costly to obtain, I do not possess the universe of earnings. I used the OPD-UHH to determine the pool of potential professional players, for whom I gathered earnings data. I gathered earnings data for the top 1% of players in terms of hours played and table-hours played.[68] In order to construct the earnings condition measure, I also gathered earnings data for a random 1% sample of all other players participating in these games.

To generate a predicted skill level for the amateur portion of the player pool, it is necessary to link the two data sets together. In order to map these 10-minute snapshots into daily labor supply variables, I do the following. Total hours worked for individual i on date d ($Hours_{id}$) is calculated as the number of snapshots the player appears in on date d, multiplied by 1/6th (because 10 minutes is 1/6th of an hour). Analogously, $Table\text{-}hours_{id} = Tables\ Played\ Simultaneously_{id} * Hours_{id}$. Merging the two datasets together, I create a panel dataset with one observation per unique player * game type * stake * date combination. Figure 14 displays the relationship between the two datasets, documenting the unique and shared variables within each. Figure 15 plots the relationship between hands played (Earnings Data) and table-hours played (OPD-UHH) at the player $\times$ date level. The clear trend and correlation of 0.94 suggests that the data are appropriately linked.

[68] All play at any stake with a minimum full buyin of at least \$100 was used for this calculation.

Figure 14: Dataset Characteristics and Matching

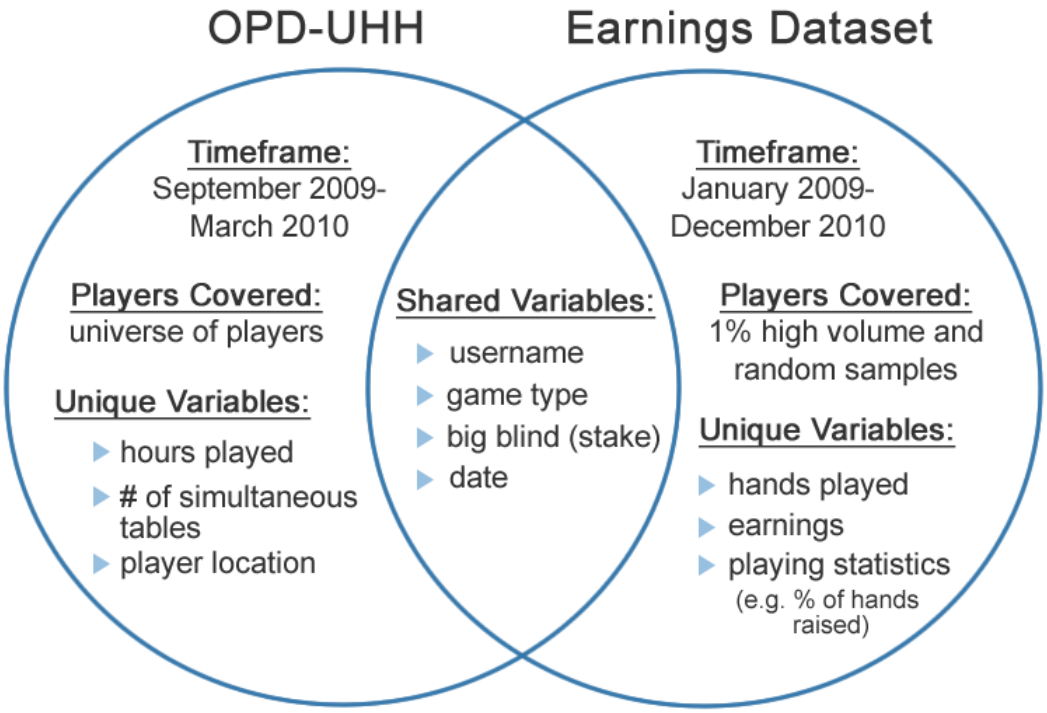
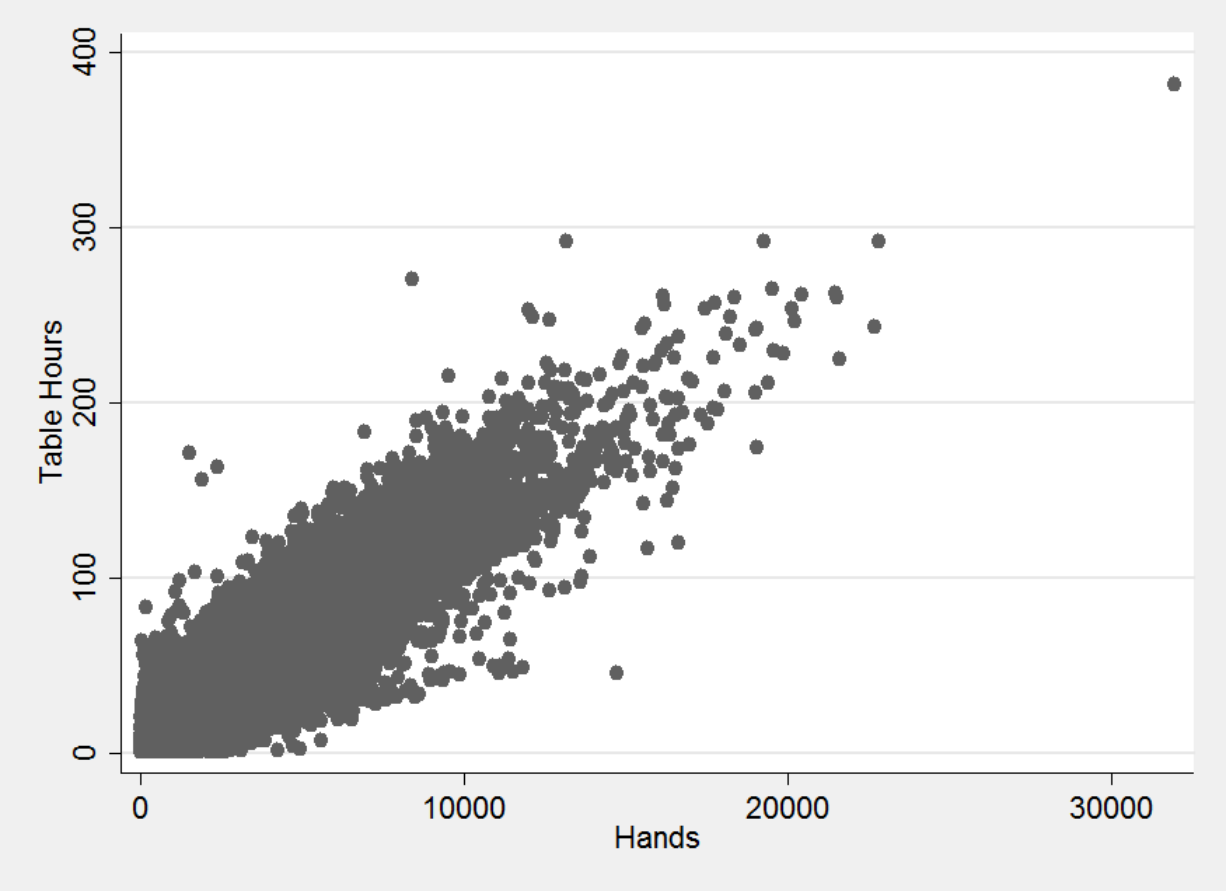


Figure 15: Dataset Alignment



Notes: Table-hours come from OPD-UHH and Hands come from the Earnings Dataset. Each observation is one player-date. Hands are a noisy function of table-hours played (roughly 65 hands per table-hour).

16.1 Defining Professionals

Throughout this analysis, I focus my attention on the labor supply behavior of "professional" players. Given that I analyze a single stake, \$1/\$2 No-Limit Texas Hold'Em 6-max, I only consider players for whom this is their most played game over the labor supply study period. I define professional players using data from the earnings dataset for the 9-month pre-period from January through September 2009. The two criteria used are: i) playing volume – a player must devote significant time to this endeavor; and ii) success – a player must be playing with the purpose of deriving income. The baseline

Table 16: Summary Statistics

	Baseline Sample	Top Pro Sample	Marginal Pro Sample
Mean Hourly Earnings	$30.49	$34.66	$22.56
	($40.47)	($35.15)	($48.40)
Mean Earnings	$6,855.61	$7,415.79	$5,785.44
	($8,171.64)	($8,604.59)	($7,213.88)
Mean Hands Played	108,190	96,448	130,623
	(71,324)	(61,502)	(83,044)
Mean Hours Played	214.2	202.5	236.6
	(83.7)	(75.47)	(94.0)
N	195	128	67

Notes: The player characteristics are averages by player over the roughly 11-week labor supply study period. Standard deviations in parentheses.

sample of professionals is defined as players that played at least 10,000 hands during the pre-period with positive earnings. A more restrictive "Top Pro" sample is defined as those who played at least 25,000 hands during the pre-period and had a win rate of at least 3 bets per 100 hands - a common benchmark within the poker community. Finally, I refer to those players who meet the first but not the second definition, as "Marginal Pros."

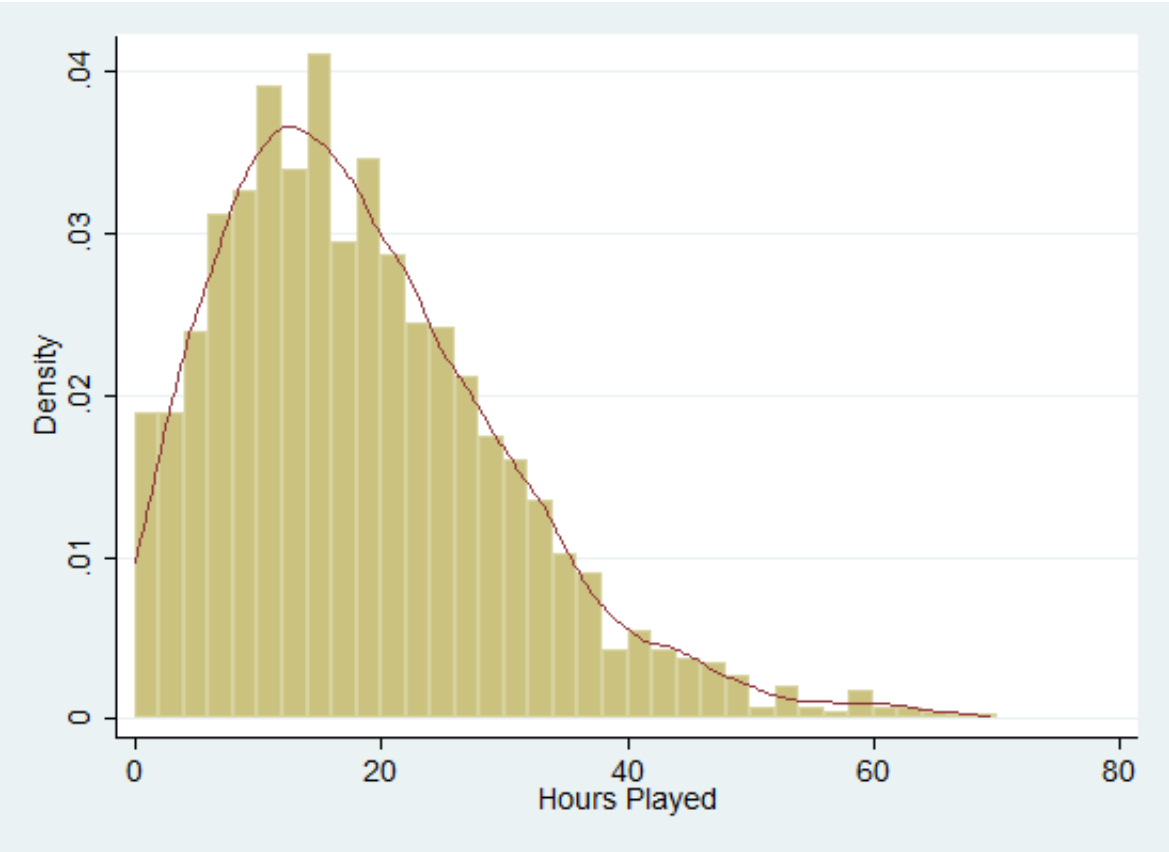
Table 16 provides descriptive statistics about playing performance and frequency for these three categories of professional players during the labor supply study period. Notably, average hourly earnings are relatively high at $30.49, significantly higher than the mean wage in the US during this period.[69] The *a priori* identified Top Pros have higher hourly earnings, $34.66, relative to the Marginal Pros $22.56, and play slightly fewer hours.

One strength of this setting relative to those used previous studies, is that poker players can literally choose any time to start and stop play. Figure 16 illustrates this

[69]Mean hourly wages in 2009 were lee than <$22.50 according to the Bureau of Labor Statistics.

phenomenon, showing that there is wide distribution in weekly hours worked, ranging from zero to over 70, with no significant clumping at 40 hours or other common thresholds.

Figure 16: Distribution of Weekly Hours Played



Note: Based on weekly observations for all 195 pros over the roughly 11-week labor supply study period.

17 Empirical Framework

When a poker player sits down at a table, virtual or otherwise, one of their first goals is to gauge the skill level and playing styles of their opponents.[70] In online poker, where many pros use software that tracks every hand they have ever played, it can be very

[70] As Matt Damon's character in the movie Rounders says, "If you can't spot the sucker in the first half hour at the table, *you* are the sucker."

easy to identify weak players. If a pro has ever played against player x, they can know within seconds of sitting down and have access to powerful statistics about player x's previous behavior. Further, Full Tilt's software allowed players to color code notes and pros often do this based on these statistics (e.g. red = good, yellow=ok, green = weak player). As such, it can be quite easy to look across all tables being played to get a feel for expected earnings. The goal of this empirical analysis is to estimate the labor supply response of professional poker players to changes in these earning conditions. Are poker players optimizing by shifting their labor supply intertemporally, or do they maintain the same schedule regardless of earning conditions?

To undertake this analysis, I adopt a three-step approach. In the first step, I estimate the skill level of all non-pro players within the sample using their observable playing characteristics. This skill level is measured in units of "expected dollar contribution per table-hour". An inexperienced amateur player would have a very large value, as they are likely to lose money quickly (e.g. contribute a lot of money to the rest of the participants). A breakeven player would have a value of 0, and a winning player would have a negative value (on average they take money from the rest of the player pool). In the second step, I aggregate this measure up to the player pool level, creating an overall amateur contribution measure for each 10-minute snapshot. In the final step, I estimate the impact of this measure on the labor supply behavior of professionals. The underlying identification assumption is that the choice of when and how much to play by amateurs is unrelated to the playing decision of pros, except through its impact on expected earnings.

17.1 Predicting "Skill" of Amateurs

In order to evaluate earning conditions for professional poker players, I need to know the skill level of their opponents (are they really bad, or just average). To this end, I estimate the predicted skill level of each non-pro player based on their previous playing history.[71] The measure used, *ExpectedContribution*, is in units of dollars lost per table-hour. A very weak player would thus have a very high *ExpectedContribution*; they lose money rapidly and hence are contributing a lot of money to the rest of the player pool. To generate these predictions, I use the 1% random sample. Losses per table-hour in the present time-period are modeled as a function of previous play.[72] As such, a players skill level can change over time as they learn and develop, or regress. For this estimation procedure, players are grouped into three categories.

Category 1: Players with earnings data (the random 1% sample) and previous playing experience - Assigned *ExpectedContribution* based on recent playing style, earnings, experience, and number of tables played.

Category 2: Players without earnings data but with previous playing experience - Assigned *ExpectedContribution* based on recent experience and number of tables played.

Category 3: New Players - Assigned *ExpectedContribution* of the average observed contribution of all first time players in the 1% sample.

The estimates from category 1 are generated from the following OLS regression framework:

$$Loss_{iw} = \beta_0 + \mathbf{X}_{iw}\mathbf{B} + \mathbf{Z}_{iw}\mathbf{C} + \varepsilon_{iw} \tag{7}$$

where Loss_{iw} is the loss per table-hour for player i in week w, $\mathbf{X}$ is a vector of individual

[71] Players who would qualify as a pro but have a different main game then \$1/\$2 NLHE 6-max are not included in the amateur group.

[72] I use the previous 3-weeks of play as a main specification, but results are qualitatively similar using 2, 4, or 6 weeks.

level controls about playing volume over the previous 3 weeks (e.g. average number of tables played simultaneously) available in the OPD-UHH and $\mathbf{Z}$ is a vector of individual controls about playing style and earnings over the previous 3 weeks available in the earnings dataset (e.g. percent of hands played). The predicted values are saved, and become the *ExpectedContribution* estimates for that player in that time period.

To generate skill estimates for category 2 players (those without earnings data), I run the modified version of equation 1 seen below on the random 1% sample:

$$Loss_{iw} = \beta_0 + \mathbf{X}_{iw}\mathbf{B} + \varepsilon_{iw} \tag{8}$$

where all variables are as defined above. I then use the regression coefficients to compute the *ExpectedContribution* values for the players outside the random 1% sample.[73] Finally, for group 3 (the new players), I simply assign the average observed contribution for all first time players in the 1% random sample.[74]

This procedure generates an *ExpectedContribution* estimate for each individual player in my sample, for every time period they are an active player. Next I translate these individual level skill measures into an overall measure for the entire amateur portion of the player pool, for each ten-minute interval, as seen below:

$$AmExpCont_t = \sum_i (ExpectedContribution_{it} * Tables_{it}) \quad \forall \ t \tag{9}$$

where $AmExpCont_{\mathrm{t}}$ is the total Amateur Expected Contribution occurring during ten-minute snapshot t. It is formed as the sum of the expected contribution per table-hour of each amateur, multiplied by the number of tables they are playing. Hence it is in

[73] Recall the category 2 players have a playing volume history ($\mathbf{X}$) but no earnings data history, hence the ommision of $\mathbf{Z}$.

[74] The regression results for both categories of players are available from the author upon request.

units of dollars per hour. Put simply, 1/6th * $AmExpCont_t$ is the average expected dollar amount that the amateur portion of the player pool is likely to contribute (i.e. lose) to the professional portion of the player pool, during that ten-minute time frame.[75]

17.2 Labor Supply Responses to Earning Conditions

The empirical model used to test how the expected contribution from amateurs affects the labor supply of professionals is:

$$\mathbf{outcomes}_{it} = \alpha_1 Ln(AmExpCont_t) + (player_i \times month_t \times DOW_t \times hour_t) + \varepsilon_{it} \quad (10)$$

where $\mathbf{outcomes}_{it} \in \{played,\ start,\ quit,\ ln(tablesplayed)\}$, $AmExpCont_t$ is as defined in the previous subsection, and $player_i \times month_t \times DOW_t \times hour_t$ is a player by month by day-of-week by hour fixed effect. Each cluster within this fixed effect contains 24 or 30 observations.[76] α_1 is the coefficient of interest, and for the binary outcomes (*played, start, quit*) it is interpreted as a 1% increase in $AmExpCont_t$ is associated with an α_1 percentage point increase in the outcome (e.g. probability of playing). When considering ln(*tablesplayed*) as the outcome, α_1 is interpreted as an elasticity.

To examine whether different groups of professionals respond differently to shocks in the player pool, I also consider the following interaction model:

$$\mathbf{outcomes}_{it} = \alpha_1 Ln(AmExpCont_t) + \alpha_2(TopPro_i \times Ln(AmExpCont_t)) \quad (11)$$

[75] 1/6th* $AmExpCont_t$ to convert dollars per hour to dollars per ten minute.

[76] There are six ten-minute intervals within each hour of the day, and a given day of the week (e.g. Tuesday) occurs either 4 or 5 times per month.

$$+(player_i \times month_t \times DOW_t \times hour_t) + \varepsilon_{it}$$

where $(TopPro_i \times Ln(AmExpCont_t))$ is an interaction of being a Top Pro, as previously defined, and the log of the amateur expected contribution. It shows any additional effect that earning conditions have on the subsample of Top Pros, relative to a reference group of Marginal Pros.

18 Results

18.1 Main Results

Table 17 shows the main results for estimating the effect of earning conditions on the labor supply decisions of professionals. The data are high-frequency panel observations, with an observation occurring at the player × 10-minute level. I compare outcomes within a specific player × hour × day-of-week × month cluster. For example, a player's decision to play at 5:40pm on Monday, Oct 6th is compared to their playing decision at all six 10-minute intervals in the 5pm hour on Mondays Oct 13th, Oct 20th, and Oct 27th, as well as the other five 10-minute intervals occurring during the 5pm hour on Monday, Oct 6th. Hence, this analysis addresses the question of whether a specific professional is more or less likely to play, during essentially the same time frame, based on prevailing earning conditions. The results suggest that professionals do respond to earning conditions, with Top Pros intertemporally shifting to better times to play, while Marginal Pros either do not intertemporally shift or shift towards worse playing times. Both types of pros adjust along the intensive margin, playing additional tables during periods with high expected earnings.

Table 17: The impact of earning conditions on whether and how much to play

	Played				Ln(*Tables*)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln(*AmExpCont*)	0.00794***	-0.00970***	-0.00985***	-0.0209***	0.0958***	0.0871***	0.0982***
	(0.00180)	(0.00300)	(0.00305)	(0.00294)	(0.0153)	(0.0248)	(0.0217)
Top Pro x Ln(*AmExpCont*)		0.0270***	0.0271***	0.0352***		0.0138	0.0117
		(0.00368)	(0.00368)	(0.00357)		(0.0309)	(0.0268)
Player x Month x DOW x Hour FE	Yes	Yes	Yes	No	Yes	Yes	No
Player x DOW x Hour FE	No	No	No	Yes	No	No	Yes
Additional Date Control	Quartic	Quartic	Weekly FEs	Quartic	Quartic	Quartic	Quartic
Observations	1,789,913	1,789,913	1,789,913	1,789,913	172,532	172,532	172,532
R-squared	0.358	0.358	0.358	0.204	0.655	0.655	0.527

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

I first focus on a discrete measure of labor supply, whether the individual is playing. The dependent variable in columns 1-4 is an indicator equal to one if the the player is observed playing. The result in column 1 suggests that pros are more likely to play when earning conditions, as proxied by ln(*AmExpCont*), are better. The point estimate of .00794 means that a 1% increase in *AmExpCont* leads to a .00794 percentage point increase in the probability that a professional is playing. To put this in context, a one standard deviation change in *AmExpCont* (within a cluster) is roughly 11%, and the mean probability of a professional playing in any time period is .0946. Thus, a one standard deviation improvement in earning conditions leads to roughly a 1% increase in the probability that a professional plays.[77]

Previous studies have demonstrated that labor supply responses can vary based on experience or skill-level of the worker (Camerer et al., 1997; Fehr and Goette, 2007; Farber, 2014). As such, in columns 2-4 I use the interaction model seen in equation 11 to test for heterogeneity in labor supply responses across professional type. The coefficient estimate on *AmExpCont* is now just the labor supply response for the reference group, the *a priori* identified Marginal Pros. Among this group, there is actually a reduction in play as earning conditions improve. The positive coefficient seen in column 1 is driven by the Top Pros, who are .017 percentage points more likely to play for every 1% increase in *AmExpCont*.[78] Moving to column 3 and changing from a quartic time trend to a more flexible time structure with week dummies leaves results relatively unchanged. Since the base fixed effects structure is asking a lot of the data, in column 4 I use an alternative fixed effects structure at the Player × Hour × DOW level. While the magnitudes change, the results are qualitatively identical. Marginal Pros are less likely to play when the Amateur Expected Contribution is higher, while Top Pros are more likely to play.

[77] A one standard deviation change in playing conditions increases probability of playing by .00794 * 11.1= .088 percentage points on a base of 9.46 percentage points.

[78] Hence, a 1 standard deviation improvement in earning conditions increases the probability of a Top Professional playing by about 2%.

In Columns 5-7, I examine behavior on the intensive margin. Given that a professional is playing, do they adjust how many tables they are playing simultaneously in response to earning conditions? The result in column 5 implies that yes, professionals do respond along this margin. This coefficient is an elasticity, suggesting that for a 1% increase in *AmExpCont*, professionals increase the number of tables played by .096%. The interaction model in column six reveals that this adjustment occurs among both categories of pros. While Top Pros might be slightly more responsive, both types of professionals add additional tables during high earning time periods. Results are qualitatively identical in column 7, with the alternative fixed effects structure.

Professional poker players do respond to earning conditions, and as a whole they do so in the manner that neoclassical theory predicts, intertemporally substituting towards labor when facing better earning conditions. However, this interpretation masks significant heterogeneity across player type. Top Pros are more likely to play during time periods with better earning conditions while marginal pros are slightly less likely to play. Conditional on playing, both types of professionals add additional tables during high earning periods. To further explore the differences between player types in the playing decision, in the next section I investigate the decision of when to start and when to stop playing.

18.2 The Start and Quit Decisions

Table 18 shows the results for estimating the effect of earning conditions on the decision of when to start and when to quit a poker session. The variable $start_{it}$ is defined as an indicator equal to 1 if the player is observed playing in time period t, but is not observed playing in period *t-1* or *t-2*.[79] Similarly, $quit_{it}$ is defined as an indicator

[79] I prefer this definition to an alternative where the player only must be absent in time period t-1, as it is less likely to count a player stopping for a few minutes for a bathroom break as starting another session. Results using this alternative definition are qualitatively identical but less precise than those

equal to 1 if the player is observed playing in time period t, but is not observed playing in period *t+1* or *t+2*. These results are estimated using both the pooled model from equation 10 and the interaction from equation 11.

using the preferred specification.

Table 18: The impact of earning conditions on when to start and quit a poker session

	Start			*Quit*			*Quit2*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln(*AmExpCont*)	0.00179**	-0.000701	-0.00242**	0.0145	0.0122	0.0271**	0.000880
	(0.000749)	(0.00126)	(0.00114)	(0.00917)	(0.0149)	(0.0120)	(0.0148)
Top Pro x Ln(*AmExpCont*)		0.00379**	0.00498***		0.00363	-0.0158	0.00618
		(0.00154)	(0.00137)		(0.0186)	(0.0148)	(0.0185)
Ln(*Session Duration*)				0.0823***	0.0823***	0.0601***	0.0808***
				(0.00121)	(0.00121)	(0.000995)	(0.00120)
Player x Month x DOW x Hour FE	Yes	Yes	No	Yes	Yes	No	Yes
Player x DOW x Hour FE	No	No	Yes	No	No	Yes	No
Additional Date Control	Quartic	Quartic	Quartic	Quartic	Quartic	Quartic	Quartic
Observations	1,600,753	1,600,753	1,600,753	172,532	172,532	172,532	172,363
R-squared	0.111	0.111	0.042	0.299	0.299	0.191	0.295

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The result in column 1 suggests that among all professionals an improvement in earning conditions leads to a small increase in the probability of starting a session. Moving to the interaction model in column 2, reveals significant heterogeneity across player types. Top Pros are significantly more likely to start playing during a time period when earning conditions are abnormally good. In contrast, Marginal Pros are unresponsive or perhaps slightly less likely to start playing during a good time period. These differences are even more stark when using the alternative fixed effects structure without a month component seen in column 3.

Columns 4-7 provide estimates for the impact of earning conditions on the decision of when to quit a session. Here, I add a control for the duration of the session, which is a strong predictor of when a player will quit. The pooled model in column 4 suggests that there is no statistically significant change in the probability of quitting a session based on earning conditions, though if anything players may be slightly more likely to quit when earning conditions are good. Moving to column 5, reveals that there is no significant difference in quitting behavior across player types. However, this result is not particularly robust, as moving to the more restrictive fixed effects model suggests that Marginal Pros are more likely to quit during time periods with better earning conditions. Finally, Column 7 uses an alternative definition for quit, where quit is set equal to 1 in the first time period that a player is unobserved, following a string of present observations. Results here are similar to those seen in column 5 and suggest that there is no strong relationship between the decision of when to quit and earning conditions.

Considering the main estimates regarding the probability of playing in conjunction with these estimates regarding the probability of starting or quitting a session provides a richer picture of the labor supply decisions of professional poker players. Top Pros are more likely to play during periods with better earning conditions, and this seems to be

driven primarily by the decision of when to start a session. Marginal Pros are slightly less likely to play during periods with better earning conditions, potentially driven by a reduction in the probability of starting a session when earning conditions are good. This asymmetry could potentially be due to differences in the tracking software and notes taken by these different groups of pros. With standard online poker tracking software and color-coded notes, a pro can easily scan the playing lobby to get a rough idea about earning conditions before they start to play. If the Marginal Pros are less likely to have these aids than Top Pros, they may not realize what earning conditions are like until ex post, when they start to play and see how opponents are behaving. This could also explain why both types of pros adjust along the intensive margin. Conditional on playing, both types of pros realize what earning conditions are, and they add additional tables when earning conditions are better.

18.3 Robustness - Addressing the Exogeneity Assumption

One concern with the analysis thus far, is that for the estimates to be causal, variation in *AmExpCont* must be exogenous within a day-of-week by hour by month cluster. However, this assumption would be violated if some unobserved factor was correlated with the playing decisions of both amateurs and pros. For example, a major winter storm could lead to an increase in play by both market segments as amateur players have a workday off, while the value of leisure time for pros is conceivably lower when outdoor activities are unavailable. While regional stories like this are mitigated by the global nature of this market, it is still possible that omitted variable bias is contributing to my result.[80] To address this concern, I add in an additional control for the volume

[80]The 195 professionals live in 28 different countries, with a maximum of 41% from any one country (USA).

of amateurs in the market.[81] Thus, the only identifying variation left in *AmExpCont* is variation in the average quality of the amateur player pool. This is cleaner variation in the sense that it is difficult to come up with a story for why average amateur quality is correlated with when professionals choose to play, other than through the channel of expected earnings.

Table 19 documents the results of this analysis for the four main outcomes. The result in column 1 reaffirms the idea that Top Pros are more likely to play when earning conditions are better. Perhaps surprisingly, this result is stronger than before as conditional on the covariates the volume of amateurs playing is actually negatively correlated with the probability a pro is playing.[82] Column 2 shows that the earning condition elasticity of tables shrinks slightly under this specification, but remains significant for both types of Pros.[83] Finally, the results in column 3 reaffirm that Top Pros are more likely to start playing when earning conditions are better, while there is little to no effect of earning conditions on the decision to quit playing. Overall, the results are qualitatively unchanged and mitigate concerns about omitted variable bias.

18.4 Aggregate Daily Measures

One question that I am unable to address with high-frequency data is what impact does *AmExpCont* have on the earnings of professionals? In other words, is *AmExpCont* a good proxy for earning conditions? Unfortunately, earnings data at the level of ten-minute intervals is not available. However, I can aggregate the $AmExpCont_t$ up from the ten-minute interval to the daily level, by simply summing all 144 measures within a

[81] This is defined as the total number of seats occupied by amateurs, hence an amateur playing two tables is counted twice.

[82] This would occur if for example, a major live poker tournament series leads to additional advertising for online poker, driving up the volume of play by amateurs, but driving down the volume of play by professionals who will substitute from online poker to participating in the tournament series.

[83] Occasionally, poker sites offer rewards for playing multiple tables at once. This could explain the positive relationship between amateur volume and number of tables played by pros.

Table 19: Robustness - Controlling for Amateur Volume

	Played (1)	*Ln(Tables)* (2)	*Start* (3)	*Quit* (4)
Ln(*AmExpCont*)	-0.00677**	0.0777***	-0.000318	0.0115
	(0.00304)	(0.0250)	(0.00127)	(0.0150)
Top Pro x Ln(*AmExpCont*)	0.0269***	0.0135	0.00379**	0.00361
	(0.00368)	(0.0309)	(0.00154)	(0.0186)
Ln(*Session Duration*)				0.0823***
				(0.00121)
Ln(*Amateur Seats*)	-0.0122***	0.0428***	-0.00163**	0.00325
	(0.00182)	(0.0147)	(0.000755)	(0.00881)
Observations	1,789,913	172,532	1,600,753	172,532
R-squared	0.358	0.655	0.111	0.299

Notes: All specifications include player x month x DOW x hour fixed effects and a quartic time trend. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

day to get a daily measure $AmExpCont_{\mathrm{d}}$.[84] Then, in the following regression framework I can look at the impact of this measure on the daily earnings of pros:

$$\mathbf{outcomes}_{id} = \alpha_1 Ln(AmExpCont_d) + (player_i \times month_d \times DOW_d) + \varepsilon_{id} \quad (12)$$

where $\mathbf{outcomes}_{id} \in \{earnings,\ table\text{-}hours\}$, and $AmExpCont_{\mathrm{d}}$ is as defined above.

The results of estimating equation 12 are seen in Table 20. Column 1 pools all pros together, and the coefficient of -49.88, means that a 1% increase in $AmExpCont_{\mathrm{d}}$ is associated with a \$49.88 decrease in daily earnings for a professional. However, this

[84] When there is a missing observation for $AmExpCont_t$, as occurs $< 1\%$ of the time when labor supply data is missing, I use a simple linear interpolation based on $AmExpCont_t$ in the surrounding time periods.

estimate is extremely noisy with a 95% confidence interval of -$362 to $262. Moving to the interaction model in Columns 2-3 does little to provide a clearer picture. Finally, Column 4 adds in the control for amateur playing volume. While this does not increase precision, it does move the point estimates to a positive $266.70 for Marginal Pros and $171.67 for Top Pros.

Table 20: Aggregated Daily Observations

	Earnings				*Table-hours*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(*AmExpCont*)	-49.88	11.48	-40.36	266.7	5.986*	4.495	-1.474	2.097
	(159.0)	(245.9)	(204.9)	(328.5)	(3.467)	(5.364)	(4.656)	(7.166)
Top Pro x Ln(*AmExpCont*)		-93.89	-82.19	-95.03		2.282	8.305	2.292
		(287.1)	(225.4)	(287.1)		(6.262)	(5.122)	(6.262)
Ln(*Amateur Seats*)				-310.5				2.918
				(265.1)				(5.781)
Player x Month x DOW FE	Yes	Yes	No	Yes	Yes	Yes	No	Yes
Player x DOW FE	No	No	Yes	No	No	No	Yes	No
Additional Date Control	Quartic	Quartic	Quartic	Quartic	Quartic	Quartic	Quartic	Quartic
Observations	13,570	13,570	13,570	13,570	13,570	13,570	13,570	13,570
R-squared	0.311	0.311	0.101	0.311	0.554	0.554	0.368	0.554

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

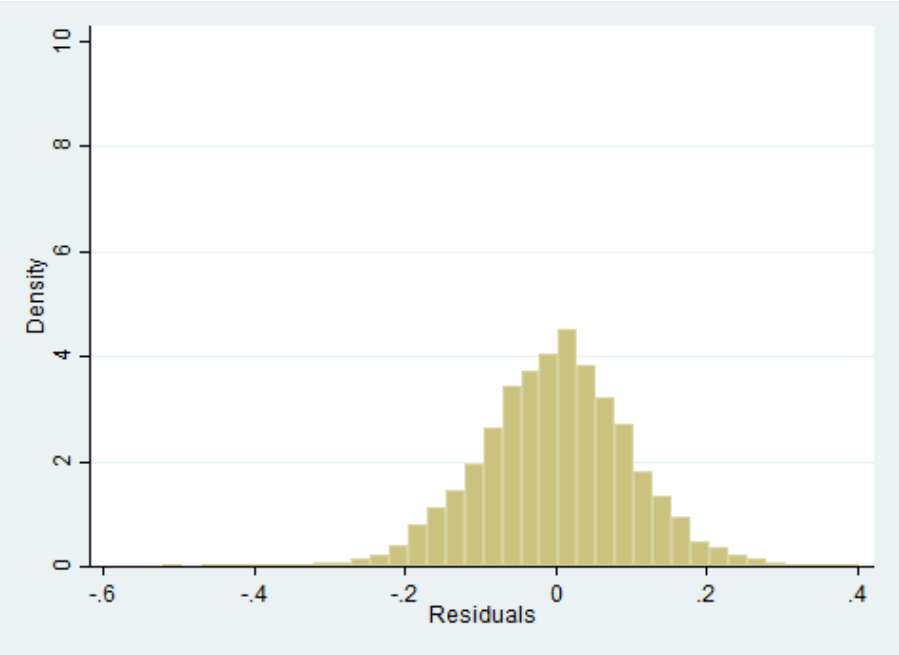
Unfortunately, the imprecision in the earnings estimates prohibit any strong conclusions from being drawn regarding the impact of $AmExpCont_{\mathrm{d}}$ on realized earnings. This noisiness is a due to at least 3 factors: 1) classical measurement error inherent in poker, as realized earnings over a short-time frame are a very noisy indicator of expected earnings; 2) by regressing a daily measure of earning conditions on realized earnings for a player that occur during only a small subset of that day, this is essentially an intent to treat effect, rather than a treatment effect;[85] and 3) there is significantly less identifying variation in the daily measure of AmExpCont relative to the high-frequency measure. This is documented in the histograms seen in Figure 17.

In columns 5-8, I examine the impact of $AmExpCont_{\mathrm{d}}$ on daily table-hours played. Although the signs are generally positive, suggesting that pros are playing slightly more on days with better earning conditions, the estimates remain quite noisy. Again, I caution against drawing any strong conclusions from this more aggregate sample.

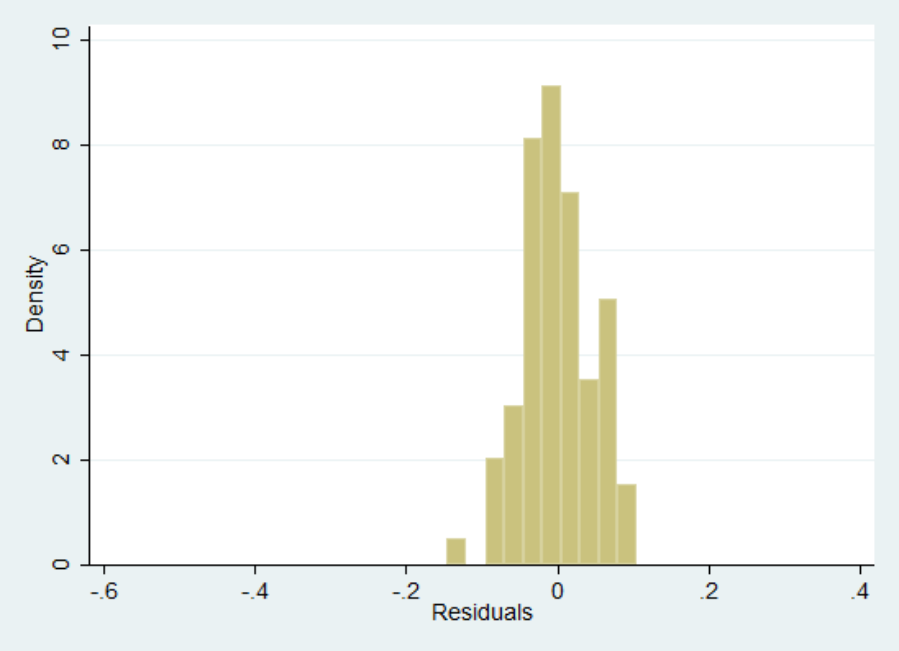
[85] This is especially problematic if earning conditions are not strongly correlated within a day.

Figure 17: Histogram of Residual Variation in *AmExpCont*

(a) High Frequency Observations



(b) Aggregated Daily Observations



Notes: Figure 17a depicts the distribution of residual variation in $AmExpCont_t$, *after regressing it on a quartic time trend and a player* $\times$ *month* $\times$ *DOW* $\times$ *hour* fixed effect. Figure 17b is the analogous figure at the daily level, using a *player* $\times$ *month* $\times$ *DOW* fixed effect.

19 Conclusion

Neoclassical models of lifecycle labor supply predict that a worker should increase hours worked when facing a transitory increase in wages. However, empirical results have been mixed with estimated elasticities ranging from -1 to 1.5. Disagreement persists in part due to the scarcity of labor markets in which workers are able to choose their labor supply on a daily basis. Even among the recent set of studies that allow for a reasonable amount of flexibility in labor supply, workers still face significant constraints (Camerer et al., 1997; Oettinger, 1999; Farber, 2005; Fehr and Goette, 2007; Crawford and Meng, 2011; Farber, 2014). I overcome this issue by focusing on a new sample of workers. Online poker players have complete flexibility in hours worked and their expected wages are changing constantly based on the overall composition of the player pool. Using high frequency labor supply data from the Online Poker Database of the University of Hamburg, I analyze the labor supply decisions of professional online poker players.

Using variation in earning conditions created by the amateur portion of the player pool, I find that the Top Pros respond according to neoclassical theory. They substitute towards labor when facing better earning conditions than are typical during that day of the week and hour. In contrast, weaker pros are actually slightly less likely to play during good times. This is potentially due to a combination of target earning (some specifications indicate that Marginal Pros are more likely to quit early when earning conditions are good) and lack of *a priori* knowledge of earning conditions (they are no more likely to start a session during good times). These results support previous findings that there is heterogeneity in labor supply responses within a particular job (Camerer et al., 1997; Fehr and Goette, 2007; Farber, 2014). In this setting, where success as a poker player depends on skills including patience and the ability to focus on expected rather than realized earnings, perhaps it is not surprising that more successful pros also appear more rational in a neoclassical sense.

Finally, this paper contributes to the growing literature examining the impact of financial incentives on effort provision of workers.[86] The study most similar to mine is Fehr and Goette (2007) who find that when facing a better earning conditions, bike messengers work additional shifts but provide less effort per shift. The primary difference between our settings is the rigidity of shift length. Since their workers must choose a lumpy number of hours associated with a discrete number of shifts of fixed length, it can be optimal to exert less effort over additional shifts. However, in my setting shift length is continuous. I find a small, but statistically significant earning condition elasticity of effort of roughly 0.1. This has implications for firms considering short-term productivity bonuses, or piece-rates, to try to meet a tight deadline.[87] My results suggest that these measures would lead to an increase in effort provision by workers.

[86]See (Prendergast, 1999) for an overview of this literature.

[87]It must be a transitory wage change, so as not to impact lifetime earnings, or their would be a potentially offsetting income effect.

References

Altonji, Joseph G. 1986. "Intertemporal substitution in labor supply: Evidence from micro data." *The Journal of Political Economy* :S176–S215.

Baker, George, Michael Gibbs, and Bengt Holmstrom. 1993. "Hierarchies and compensation: A case study." *European Economic Review* 37 (2):366–378.

———. 1994a. "The internal economics of the firm: evidence from personnel data." *The Quarterly Journal of Economics* :881–919.

———. 1994b. "The wage policy of a firm." *The Quarterly Journal of Economics* :921–955.

Barnes, Christopher M and David T Wagner. 2009. "Changing to daylight saving time cuts into sleep and increases workplace injuries." *Journal of Applied Psychology* 94 (5):1305.

Bognanno, Michael L. 2001. "Corporate tournaments." *Journal of Labor Economics* 19 (2):290–315.

Brick, Ivan E, Oded Palmon, and John K Wald. 2012. "Too much pay-performance sensitivity?" *Review of Economics and Statistics* 94 (1):287–303.

Broughton, J, M Hazelton, and M Stone. 1999. "Influence of light level on the incidence of road casualties and the predicted effect of changing 'summertime'." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 162 (2):137–175.

Browning, Martin, Angus Deaton, and Margaret Irish. 1985. "A profitable approach to labor supply and commodity demands over the life-cycle." *Econometrica* :503–543.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. 1997. "Labor supply of New York City cabdrivers: One day at a time." *The Quarterly Journal of Economics* 112 (2):407–441.

CDC. 2005-2010. "10 Leading Causes of Injury Deaths by Age Group Highlighting Unintentional Injury Deaths, United States."

Coren, Stanley. 1996. "Daylight savings time and traffic accidents." *New England Journal of Medicine* 334 (14):924–925.

Crawford, Vincent P and Juanjuan Meng. 2011. "New york city cab drivers labor supply revisited: Reference-dependent preferences with rational expectations targets for hours and income." *The American Economic Review* 101 (5):1912–1932.

Dickens, William T and Shelly J Lundberg. 1993. "Hours Restrictions and Labor Supply." *International Economic Review* :169–192.

Dohmen, Thomas and Armin Falk. 2011. "Performance pay and multidimensional sorting: Productivity, preferences, and gender." *American Economic Review* :556–590.

Doleac, Jennifer L and Nicholas J Sanders. 2013. "Under the Cover of Darkness: How Ambient Light Influences Criminal Activity." Working Paper.

Ehrenberg, Ronald G and Michael L Bognanno. 1990a. "Do Tournaments Have Incentive Effects?" *Journal of Political Economy* :1307–1324.

———. 1990b. "The incentive effects of tournaments revisited: Evidence from the European PGA tour." *Industrial and Labor Relations Review* :74S–88S.

Eriksson, Tor. 1999. "Executive compensation and tournament theory: Empirical tests on Danish data." *Journal of Labor Economics* 17 (2):262–280.

Farber, Henry S. 2005. "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers." *Journal of Political Economy* 113 (1).

———. 2014. "Why You Can't Find a Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers." Tech. rep., National Bureau of Economic Research.

Fehr, Ernst and Lorenz Goette. 2007. "Do workers work more if wages are high? Evidence from a randomized field experiment." *The American Economic Review* :298–317.

Ferguson, Susan A, David F Preusser, Adrian K Lund, Paul L Zador, and Robert G Ulmer. 1995. "Daylight saving time and motor vehicle crashes: the reduction in pedestrian and vehicle occupant fatalities." *American Journal of Public Health* 85 (1):92–95.

Fernie, Sue and David Metcalf. 1999. "Its not what you pay its the way that you pay it and thats what gets results: Jockeys pay and performance." *Labour* 13 (2):385–411.

Fridstrom, Lasse, Jan Ifver, Siv Ingebrigtsen, Risto Kulmala, and Lars Krogsgard Thomsen. 1995. "Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts." *Accident Analysis & Prevention* 27 (1):1–20.

Gelman, Andrew and Guido Imbens. 2014. "Why High-order Polynomials Should not be Used in Regression Discontinuity Designs." Tech. rep., National Bureau of Economic Research.

Gittleman, Maury and Brooks Pierce. 2013. "How Prevalent is Performance-Related Pay in the United States? Current Incidence and Recent Trends." *National Institute Economic Review* 226 (1):R4–R16.

Harrison, Yvonne. 2013. "The impact of daylight saving time on sleep and related behaviours." *Sleep Medicine Reviews* 17:285–292.

Hudson, George Vernon. 1895. "On seasonal time-adjustment in countries south of latitude 30." *Transactions and Proceedings of the New Zealand Institute* 28:734.

Imbens, Guido and Karthik Kalyanaraman. 2012. "Optimal bandwidth choice for the regression discontinuity estimator." *The Review of Economic Studies* 79 (3):933–959.

Imbens, Guido W and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics* 142 (2):615–635.

Janszky, Imre, Staffan Ahnve, Rickard Ljung, Kenneth J Mukamal, Shiva Gautam, Lars Wallentin, and Ulf Stenestrand. 2012. "Daylight saving time shifts and incidence of acute myocardial infarction–Swedish Register of Information and Knowledge About Swedish Heart Intensive Care Admissions (RIKS-HIA)." *Sleep Medicine* 13 (3):237–242.

Kamstra, Mark J, Lisa A Kramer, and Maurice D Levi. 2000. "Losing sleep at the market: The daylight saving anomaly." *The American Economic Review* 90 (4):1005–1011.

Kellogg, Ryan and Hendrik Wolff. 2008. "Daylight time and energy: Evidence from an Australian experiment." *Journal of Environmental Economics and Management* 56 (3):207–220.

Kniesner, Thomas J, W Kip Viscusi, Christopher Woock, and James P Ziliak. 2012. "The value of a statistical life: Evidence from panel data." *Review of Economics and Statistics* 94 (1):74–87.

Knutson, Kristen L, Eve Van Cauter, Paul J Rathouz, Thomas DeLeire, and Diane S Lauderdale. 2010. "Trends in the prevalence of short sleepers in the USA: 1975–2006." *Sleep* 33 (1):37.

Kotchen, Matthew J and Laura E Grant. 2011. "Does daylight saving time save energy? Evidence from a natural experiment in Indiana." *Review of Economics and Statistics* 93 (4):1172–1185.

Lahti, Tuuli, Esa Nysten, Jari Haukka, Pekka Sulander, and Timo Partonen. 2010. "Daylight saving time transitions and road traffic accidents." *Journal of Environmental and Public Health* 2010.

Lazear, Edward P. 1992. "The job as a concept." *Performance measurement, evaluation, and incentives* :183–215.

———. 2000. "Performance Pay and Productivity." *American Economic Review* 90 (5):1346–1361.

Lazear, Edward P and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* :841–864.

Legree, Peter J, Tonia S Heffner, Joseph Psotka, Daniel E Martin, and Gina J Medsker. 2003. "Traffic crash involvement: Experiential driving knowledge and stressful contextual antecedents." *Journal of Applied Psychology* 88 (1):15.

Levitt, Steven D and Thomas J Miles. 2014. "The Role of Skill Versus Luck in Poker: Evidence From the World Series of Poker." *Journal of Sports Economics* 15 (1):31–44.

Levitt, Steven D, Thomas J Miles, and Andrew M Rosenfield. 2012. "Is Texas Hold'Em a Game of Chance-A Legal and Economic Analysis." *Georgetown Law Journal* 101:581.

Lockley, Steven W, Laura K Barger, Najib T Ayas, Jeffrey M Rothschild, Charles A Czeisler, Christopher P Landrigan et al. 2007. "Effects of health care provider work hours and sleep deprivation on safety and performance." *Joint Commission Journal on Quality and Patient Safety* 33 (Supplement 1):7–18.

Lucas, Robert E and Leonard A Rapping. 1969. "Real wages, employment, and inflation." *The Journal of Political Economy* :721–754.

Ludwig, Jens and Douglas L Miller. 2007. "Does Head Start improve children's life chances? Evidence from a regression discontinuity design." *The Quarterly Journal of Economics* 122 (1):159–208.

MaCurdy, Thomas E. 1981. "An Empirical Model of Labor Supply in a Life-Cycle Setting." *The Journal of Political Economy* 89 (6):1059–1085.

Nagin, Daniel S, James B Rebitzer, Seth Sanders, and Lowell J Taylor. 2002. "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment." *American Economic Review* :850–873.

Oettinger, Gerald S. 1999. "An Empirical Analysis of the daily Labor supply of Stadium Vendors." *The Journal of Political Economy* 107 (2):360–392.

Olders, Henry. 2003. "Average sunrise time predicts depression prevalence." *Journal of Psychosomatic Research* 55 (2):99–105.

Paarsch, Harry J and Bruce S Shearer. 1999. "The response of worker effort to piece rates: evidence from the British Columbia tree-planting industry." *Journal of Human Resources* :643–667.

Prendergast, Canice. 1999. "The provision of incentives in firms." *Journal of Economic Literature* :7–63.

Prerau, David. 2005. *Seize the Daylight: The Curious and Contentious Story of Daylight Saving Time.* Thunder's Mouth Press.

Sexton, Alison L and Timothy KM Beatty. 2014. "Behavioral responses to Daylight Savings Time." *Journal of Economic Behavior & Organization* .

Smith, Michael E, Linda K McEvoy, and Alan Gevins. 2002. "The impact of moderate sleep loss on neurophysiologic signals during working-memory task performance." *Sleep* 25 (7):784.

Sood, Neeraj and Arkadipta Ghosh. 2007. "The short and long run effects of daylight saving time on fatal automobile crashes." *The BE Journal of Economic Analysis & Policy* 7 (1).

Sullivan, John M and Michael J Flannagan. 2002. "The role of ambient light level in fatal crashes: inferences from daylight saving time transitions." *Accident Analysis & Prevention* 34 (4):487–498.

Tosi, Henry L, Steve Werner, Jeffrey P Katz, and Luis R Gomez-Mejia. 2000. "How much does performance matter? A meta-analysis of CEO pay studies." *Journal of Management* 26 (2):301–339.

Valdez, P, C Ramirez, A Garcia, and E Garcia. 1997. "Adjustment of sleep to daylight saving time during weekdays and weekends." *Chronobiology International* 14.

Varughese, Jason and Richard P Allen. 2001. "Fatal accidents following changes in daylight savings time: the American experience." *Sleep Medicine* 2 (1):31–36.

Wagner, David T, Christopher M Barnes, Vivien KG Lim, and D Lance Ferris. 2012. "Lost sleep and cyberloafing: Evidence from the laboratory and a daylight saving time quasi-experiment." *Journal of Applied Psychology* 97 (5):1068.

Winston, Clifford. 2013. "On the Performance of the US Transportation System: Caution Ahead." *Journal of Economic Literature* 51 (3):773–824.

Wolff, Hendrik and Momoe Makino. 2013. "Does Daylight Saving Time Burn Fat? Time Allocation with Continuous Activities." Working Paper.