Three Essays on Applied Microeconomics: Consumer and Employee Behaviour

by

Jie Yang

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Economics

University of Alberta

Abstract

This dissertation consists of three chapters, each of which can be considered an independent essay. The three essays contribute to labor economics, sports economics, and behavioral economics.

The first chapter focuses on pay and performance in a team environment. Working in teams increases productivity but also generates incentives to shirk. Recent research suggests that peer enforcement, coupled with financial incentives in a setting of repeated interactions, can play a role in deterring shirking in teams. This paper, entitled "*Peer Enforcement in Teams: Evidence from High-Skill Professional Workers with Repeated Interactions*," analyzes 10 years of performance and compensation data for NFL offensive linemen, a high-skill, high-salary, repeatedly interacting team, using the Hausman-Taylor estimator to control for unobservable individual-specific heterogeneity. We find evidence that teammates' effort signals reduce the salaries of individual offensive linemen, providing a low powered sanctioning mechanism for individual workers in this setting. A separate, independently monitored individual effort signal also reduces salaries.

The second chapter of my dissertation is "*Consumption Commitments and Simultaneous Insuring and Gambling: Evidence from Canada*." This paper extends recent work by Chetty and Szeidl (2007) on a classic economic research question: why do some individuals simultaneously buy insurance and gamble? This behavior is "contradictory" to von Neumann-Morgenstern expected utility theory because insurance purchase indicates risk aversion while gambling indicates risk loving. Chetty and Szeidl (2007) propose a novel explanation based on consumption commitments which

magnify risk aversion, inducing Friedman-Savage local non-concavity in the utility function. Theoretically, the paper shows that commitments increase risk loving over gambles with large uncertain payoffs but for gambles with moderate-to-large uncertain payoffs, moderate commitments amplify risk loving, while large commitments mitigate risk loving. Empirically, patterns in household decisions to participate in government-run lotteries, small prize gambles (including casino gambling, slot machines, and video lottery terminals), and to purchase life insurance support these predictions; households with large consumption commitments are more likely to participate in activities with large uncertain payoffs.

The last chapter, "*The Relationship Between Consumer Spending on Exercise, Sports Betting and Attending Sporting Events*" also utilizes data from the SHS, but a large sample - more than 145,000 households. We investigate the relationship between consumer spending on three alternative leisure time activities: sports betting, exercise and attending live sporting events. Recent proposed changes in legal sports betting, and claims about attending games and participation in physical activity motivate analysis of these categories of consumer spending. Using several versions of an Almost Ideal Demand System (AIDS) and the related Quadratic AIDS (QAIDS) models, we estimate the parameters that determine the relationship between consumer spending on these activities. Results show that betting and attending games are complements. Betting and exercise, and attending games and exercise are substitutes.

Preface

I collaborated with my supervisors, Professors Jane Ruseski and Brad Humphreys, on some of the research for this thesis. Chapter 1 was coauthored with Professor Humphreys. I was responsible for the data collection and cleaning, theoretical and empirical approach developments and econometric analysis as well as the manuscript composition. Professor Humphreys was involved in motivation, research methodology development and empirical analysis as well as manuscript composition and edit. Chapter 3 was coauthored with Professor Humphreys and Professor Ruseski. I was responsible for the data analysis, empirical methodology development, econometric analysis and the manuscript composition. Professor Humphreys and Professor Ruseski contributed to motivation, methodology development, result analysis and manuscript composition and edit.

Acknowledgements

I am indebted to my supervisors, Professors Jane Ruseski, Brad Humphreys and Haifang Huang, for their great supports throughout my PhD studies, and for their invaluable comments and suggestions on my dissertation research. I also want to thank Professors Ruseski and Humphreys for their arranging and providing data access and financial assistance to help me completing my degree.

It was my fortune to have met my supervisors. Their continual encouragement and enlightenment, and kind attentions to my education and career are essential to my PhD studies. They are the best supervisors I have ever met. I am very grateful for all I have received these years, which has led me where I am now.

I thank very much Professors Vera Brencic, Marvin Washington and Craig Depken II for accepting to serve as my dissertation committee members, for their times, and for valuable suggestions.

Very helpful discussions with Professors David Ryan, Dennis Young, Rob Simmons, Dennis Coates and David Forrest are sincerely acknowledged. I thank very much for their great comments and suggestions, and for answering questions about my research. I also want to thank Professors Valentina Galvani, Tilman Klumpp, Xuejuan Su, Landon Stuart, Constance Smith, Sebastian Fossati, Corinne Langinier, Douglas West, Heather Eckert, Andrew Eckert, Li Zhou and all other professors and colleagues who provided their insightful comments and suggestions on my thesis.

I also want to thank Professor Haifang Huang, Professor Valentina Galvani, Dr. Tomas Nilsson, and Professor Yingfeng Xu for their supports and helps during my studies at University of Alberta. I appreciate their attentions to my education and life. The Christmas and New Year parties held by them helped me enjoying my overseas studies more and becoming less homesick. In addition, I am grateful to Professor Galvani and Dr. Nilsson for their consideration for my career, for arranging a summer internship for me, and for their helpful advices.

I am very thankful to all the professors who have taught me, for their time spent on my education, and for their kind helps during my doctoral studies. I thank all faculty members and staff from the Department of Economics for all their helps and supports. In particular, I am grateful to Department Chairs, Professors Douglas West and Constance Smith, Associate Chairs, Professors Brad Humphreys, David Ryan, Todd Smith and Rick Szostak, and Ms. Audrey Jackson, Ms. Brenda Carrier, Ms. Charlene Hill and Ms. Chelsi Hudson. Their excellent work and kind helps make me feel well supported and protected all the time.

The financial support of Interfaculty Graduate Studentships provided jointly by Faculty of Medicine and Dentistry and Faculty of Arts, and Assistantships provided by the Department of Economics is gratefully acknowledged.

The analyses presented in the second and third chapters of my dissertation were conducted at the Research Data Centre (RDC), University of Alberta, which is part of the Canadian Research Data Centre Network. Those research was supported by funds to the Canadian Research Data Centre Network (CRDCN) from the Social Sciences and Humanities Research Council (SSHRC), the Canadian Institute for Health Research (CIHR), the Canadian Foundation for Innovation (CFI), and Statistics Canada. Although the research and analysis are based on data from Statistics Canada, the opinions expressed do not represent the views of Statistics Canada. I also want to thank Ms. Irene Wong for her kind help and great support in the RDC.

I finish with a final gratitude for my family. I appreciate their all the time support and understanding. Without that, I could not finish anything; and it is my family that keeps me going and going...

Contents

List of Tables

List of Figures

Chapter 1

Peer Enforcement in Teams: Evidence from High-Skill Professional Workers with Repeated Interactions

1.1 Introduction

Firms often organize employees into teams. Economists posit that firms organize workers into teams to take advantage of complementarities, increasing the productivity and output of teams beyond that of individual workers. Organizing workers into teams also generates problems. Holmstrom (1982) showed that groups of workers with inter-related productive inputs generates a moral hazard problem, in the form of an incentive for some team members to supply less effort, or shirk. Successful teams exploit complementarities and reduce the incentive for individual workers to shirk.

The incentive to shirk in teams can be mitigated in a number of ways. Che and Yoo (2001) analyzed incentives in teams when the team members repeatedly interact and observe the behavior of other team mates. Their model predicts that compensation schemes that reward an employee when co-workers perform well, and punish an employee when co-workers preform poorly, have desirable properties in a workplace featuring repeated interaction among teammates. Ishida (2006) generalize this model to the case where compensation is relative, and not absolute, and shows that similar predictions emerge from the model. The key feature in these models is the existence of sanctions imposed by team members for past behavior that can be used to reduce shirking. Kandel and Lazear (1992) make a similar point, based on a model of partnerships with peer-pressure. They point out that in order for sanctions to reduce the incentive to shirk in a team, an individual team member's "effort must affect the well-being of the rest of the team for them to have incentive" to sanction another teammate.

We develop evidence from the field that incentives like those described in the model developed by Che and Yoo (2001) exist. Mas and Moretti (2009), Ichino and Maggi (2000), and Chan et al. (2014) develop evidence of peer interaction in teams using data from cashiers in a national supermarket chain, an Italian bank, and cosmetic sales in a department store in China, respectively. Mas and Moretti (2009) focus on peer interaction and productivity; Ichino and Maggi (2000) focus on the incidence of shirking. Depken and Haglund (2011) also develop evidence of peer effects in team sports, although this evidence is based on observed productivity, not earnings of team members. We analyze the performance and compensation of members of a high-skill, high-salary team that interacts repeatedly, offensive linemen in the National Football League (NFL). Offensive line play in the NFL is complex, and the actions of individual linemen are highly inter-related. These high-skill workers interact hundreds of times over the course of a season, much like the infinitely repeated game in the model developed by Che and Yoo (2001).

We find that, after controlling for unobservable worker-specific heterogeneity, experience, and other personal and team-specific characteristics that affect performance and compensation, a specific observable effort signal generated by team mates reduces the salaries of individual offensive linemen, but has no effect on the salaries of the team mates, providing an optimal, low powered sanctioning mechanism for individual workers. This supports the predictions from the model developed by Che and Yoo (2001). We also find that another observable signal related to shirking that is monitored by independent evaluators affects only the shirking employee, suggesting that NFL teams also use standard financial incentives to deter shirking in this setting.

1.2 Motivation and Context

Che and Yoo (2001) developed a model of team behavior that includes long-term interaction among team members; the model features an infinitely repeated game in which team members observe the effort of other workers on the team and decide how much effort to supply in each period. Ishida (2006) extended this model to the case where only relative performance analysis takes place. Each workers' strategy is a function that maps all possible past effort decisions into a probability distribution over current effort decisions by team members. In this model a firm hires two identical workers to repeatedly perform a specific project or task. Each worker makes a binary effort decision k to either supply effort ("work," $k = 1$), or not supply effort ("shirk," $k = 0$) in each period. Effort requires a cost e but shirking entails no effort-related cost. The key feature of the model is that the workers experience close interaction in each period, so they perfectly observe each others effort decision. The model includes mutual monitoring, an important feature of any team-based work arrangement. The workers interact only through their effort decisions in this model, they cannot exchange side payments. The model predicts that effort decisions generate self-enforcing incentives that take the form of punishing other team members by shirking.

Firms do not observe individual effort decisions made by workers in this model. Instead, firms receive a binary signal, x_i, that is either good ($x_i = 1$) or bad ($x_i = 0$). This imperfect signal about the workers' effort decision reflects both the individual's effort and a random environmental shock that affects both workers. This common environmental shock is either favorable or unfavorable, and the probability of a favorable common shock is σ. In an alternative version of the model, the firm receives a single imperfect signal that is the result of the team effort and the environmental shock instead of a signal for each team member. The predictions of the model do not depend critically on the nature of this signal.

The team production arrangement is open ended and terminates with probability $1 - \delta$ at the end of each period, so δ indicates how long teams remain together. The workers always have incentive to shirk since effort is costly but shirking is not easily observed given the imperfect effort signal. The firm wants to induce workers to supply positive effort in each period, because the outcome when both workers provide positive effort is more valuable than other outcomes. The firm's problem is to motivate both workers to provide positive effort in every period at minimum cost using some wage scheme. Specifically, the wages cannot be negative and depend on the verifiable signals received by the firm. Clearly, an important feature of the wage scheme designed by the firm in this context is how an individual's compensation relates to the performance of other team members.

Two possible wage schemes exist in this setting: a wage scheme based on relative performance evaluation (RPE) under which a worker is penalized when other team members perform well and rewarded when other team members perform poorly; and a wage scheme based on joint performance evaluation (JPE) under which a worker is rewarded when other team members perform well and penalized when other team members perform poorly. Tournament theory represents a special form of RPE; under RPE, the worker perceived to perform the best, based on the signals received, earns the highest compensation, the second best the second highest, and so on. RPE wage schemes appear to be ineffective when workers interact closely with each other, perhaps because it leads to competition among team members, especially when workers can sabotage others in the the competition [Lazear (1989)]. JPE wage schemes overcome these problems in settings where workers interact repeatedly by reducing the negative aspects of competition while providing team members with a mechanism to deter shirking. Under JPE, a worker can shirk in order to punish another team member because the other team member's compensation will be reduced by this action. Che and Yoo (2001) show that the firm can make a larger profit under JPE than under RPE when workers repeatedly interact, because the total cost of the wage scheme is lower under JPE than RPE.

In Che and Yoo (2001)'s model, the firm tries to minimize the cost of inducing each team member to work,

$$\min \pi(1,1;w),$$

subject to incentive constraint (IC):

$$\pi(1,1;w) - e \geq (1-\delta)\pi(0,1;w) + \delta \min\{\pi(0,0;w), \pi(0,1;w)\}.$$

Under a JPE scheme, $\pi(0,1;w^J) > \pi(0,0;w^J)$, then $\min\{\pi(0,0;w^J), \pi(0,1;w^J)\} = \pi(0,0;w^J)$. Since IC is binding at w^J,

$$\begin{aligned} \pi(1,1;w^J) - e = (1-\delta)\pi(0,1;w^J) + \delta\pi(0,0;w^J) &> \pi(0,0;w^J); \\ &< \pi(0,1;w^J). \end{aligned}$$

Because $\pi(k,l;w^J)$ is supermodular in (k,l),

$$\pi(1,1;w^J) + \pi(0,0;w^J) - \pi(0,1;w^J) - \pi(1,0;w^J) > 0.$$

It follows that

$$\pi(1,1;w^J) - e + \pi(0,0;w^J) - \pi(0,1;w^J) > \pi(1,0;w^J) - e.$$

Therefore,

$$\pi(0,0;w^J) > \pi(1,0;w^J) - e.$$

In other words, a worker will shirk to punish another team member after observing his shirking action. Consequently, (shirk, shirk) is a stage-game Nash equilibrium, but a worse outcome than (work, work). Thus, given this threat of punishment, repeated play develops a subgame perfect equilibrium, (work, work).

On the contrary, under a RPE scheme, $\pi(0,1;w^R) < \pi(0,0;w^R)$. Since IC is binding at w^R,

$$\pi(1,1;w^R) - e = (1-\delta)\pi(0,1;w^R) + \delta\pi(0,1;w^R) = \pi(0,1;w^R) < \pi(0,0;w^R).$$

That is, RPE cannot implement repeated play of (work, work) as a subgame perfect equilibrium like JPE. Moreover, the cost of inducing a team member to work rather than shirk is $\pi(1,1;w^R) - \pi(0,1;w^R) = e$ under RPE, larger than $\pi(1,1;w^J) - \pi(0,1;w^J) < e$ under JPE.

Che and Yoo (2001) demonstrate that JPE wage schemes are optimal relative to RPE wage schemes in settings where workers on a team interact repeatedly, no matter how long workers remain in teams. RPE wage schemes are more likely to be optimal as the probability of favorable common shocks increases. JPE wages schemes have another important property in this model: the explicit incentives, those that affect wages directly, must be relatively low powered for shirking to be a credible punishment strategy. In other words, for shirking to be deterred, the deterrence mechanism must affect the other worker more than the worker who seeks to deter shirking; the effect must be asymmetric.

The presence of JPE or RPE wage schemes in a team can be empirically investigated. Given data on the earnings of individual workers who are organized into teams and interact repeatedly in production, and proxy variables for positive or negative effort signals received by firms, under RPE wage schemes, a larger number of negative signals from other workers on the team would be associated with higher earnings by a given worker, and a larger number of positive signals from other workers would be associated with lower earnings for a given worker. Under JPE wage schemes, a larger number of negative effort signals from other workers represent peer sanctions. In addition, the model developed by Che and Yoo (2001) predicts that JPE wage schemes would be more likely to be observed in a setting where the teams are relatively long-lived and where the probability of common positive shocks to team members is relatively low.

One setting where JPE wage schemes might be present is among offensive linemen in professional football leagues like the NFL. Note that we interpret the offensive line as group of workers (football players) organized into a "team" within a professional football team. Che and Yoo (2001) identify three characteristics of a team in the context of their model: (1) frequent and consistent interaction over a long period of time, (2) autonomy to make independent decisions on assignments and problem solving, and (3) encouragement to monitor and motivate each other. Offensive linemen must work together closely in a highly coordinated way in order to create running room for ball carriers when the team runs the ball and to protect the quarterback when the team passes the ball. The assignments for individual linemen are highly inter-related and complex. NFL teams attempt to reduce turnover in personnel on the offensive line, and since NFL players only gain free agency after three years of experience in the league, offensive linemen often play together for a number of seasons. During the game, the offensive line makes its own play calls at the line of scrimmage after the quarterback calls the offensive play that will be run in the huddle, suggesting that the offensive line has some independent decision making power. And anyone who has watched an NFL game would agree that offensive linemen continually monitor and motivate each other.

On an NFL team, the offensive line typically contains five players, or in this context, team members. The center plays in the middle of the line and snaps the football, handing or throwing it between his legs to the quarterback to initiate each play. Next to the center are the right and left guard, and outside the guards are the right and left tackle. Offensive linemen cannot touch the ball in most circumstances; the goal of this team is to provide open space for other players advancing the ball by running and to protect the quarterback when he passes the ball. The primary activity performed by offensive linemen is called "blocking." This activity involves physical interaction with defensive players and a large body of rules specify exactly how an offensive lineman can legally touch a defensive player and also specify illegal forms of interaction that, if detected by an official, will result in a penalty in the form of moving the ball backward by a specific distance, usually five or ten yards. Offensive linemen also must, by rule, remain absolutely still until the center snaps the ball to the quarterback and cannot run down field on passing plays.

In the context of effort supply, we interpret penalties committed by offensive linemen as a signal of shirking. Seven officials monitor play in the NFL. Three of these officials monitor offensive linemen for rule violations, as these workers are subject to a large number of rules that proscribe many specific activities. Offensive linemen can commit a number of infractions that result in a penalty if observed by an official. Two common penalties committed by offensive linemen are false starts and holding. A false start occurs when an offensive lineman moves before the center has snapped the ball to the quarterback and results in a five yard penalty. Holding occurs when an offensive lineman (or other offensive player) grabs or tackles a defensive player in a way prohibited by the rules and results in a ten yard penalty. When an offensive lineman commits holding, he supplies less effort than he would do if no penalty was committed. A false start is a mental error, but can also be interpreted as supplying less effort, in this case effort to remember the signal for the snap of the ball.

In the model developed by Che and Yoo (2001), firms receive a signal that reflects both effort supplied by workers and a common random shock. NFL teams receive signals about the effort supplied by offensive linemen. Clearly, NFL players could engage in a large number of types of shirking. In addition to the penalties discussed above, we also consider a second type of signal about effort supplied by offensive linemen, giving up sacks. On a passing play, the goal of the offensive line is to keep the defensive players from tackling the quarterback before he can throw the ball down field. A sack occurs when a defensive player tackles the quarterback in the backfield before he can throw the ball. This results in a loss of yardage, a bad outcome for the offense. We interpret a sack allowed by an offensive lineman as a signal of shirking. Implicitly, if an offensive lineman would have supplied more effort, a sack could have been avoided on a play. Unlike penalties, which are monitored by the officials, coaches and other workers must monitor offensive linemen on passing plays to determine the amount of effort supplied.

Since the effort signal includes a random shock component, shirking can be somewhat masked by ability, random events, or the presence of a strong opponent. Shirking usually occurs when player is facing an evenly-matched or stronger opponent (a negative shock), since shirking would be easily observed if facing weaker opponent (a positive shock). In practice, shirking may be determined by a coach after reviewing the game video, which is costly and inefficient. It also cannot be easily determined during a game, when coaches must make personnel decisions based on the game situation and specific player match-ups. Further, some sacks can't be awarded to specific players even after video review. But offensive linemen on the field during the game have better information about effort supply. Because of repeated interactions, they know more than anyone else about their colleagues' effort supply. So mutual monitoring and punishment would be an efficient and less costly way to deter shirking during games.

We empirically analyze the relationship between these two effort signals and the salaries earned by these players. Again, we assume that penalties committed and sacks allowed can be interpreted as signals of effort supplied by offensive linemen. Since we observe both signals received by the firm, and the salaries of the team members, this information can be exploited to determine if a JPE wage scheme exists for these high-skilled workers. Again, under JPE wage schemes, workers are penalized for poor performance by team mates, a mechanism through which workers, over the course of repeated interaction, punish other workers for supplying lower levels of effort. If penalties committed, or sacks allowed by team mates reduces the compensation of offensive linemen, then outcomes in this setting are consistent with JPE wage schemes, supporting the model developed by Che and Yoo (2001).

1.3 Data Description

Our data include information about the performance and earnings of NFL offensive linemen, and their teams, over the 2000 to 2009 regular seasons. Our basic unit of observation is an individual NFL offensive lineman over a season. Offensive line play is highly inter-related. Unlike many other positions on an NFL team, the offensive linemen must cooperate and work as a unit in order to perform well. The data set we construct contains five types of information about offensive linemen: (1) individual performance data for each season, (2) player characteristics, (3) salary, (4) financial data for NFL teams and (5) team offensive performance. The data were collected from a number of sources. The player performance and characteristics data are from several sports data websites, including the official NFL website, Stats Incorporated's fee-based STATSPASS database, and www.pro-football-reference.com; the salary data are from the USA Today NFL Salary database; the team-specific financial data are from the estimates published in *Forbes* magazine and on their website; and the team's offensive performance data are generated from data on www.pro-football-reference.com. There were 2652 player-year observations including data from 688 unique players over the 2000 to 2009 seasons. Summary statistics for the player characteristics and salary variables are shown on Table 1.1.

NFL player compensation consists of guaranteed pay in the form of a "signing bonus", an annual "base salary" paid to players who remain on the team's roster, and performance-based pay that depends on team performance and the number of plays each player participated in over the course of a season.[1] Neckermann et al. (2014) analyze such bonus pay in a different setting. Our measure of salary is the "cap value" accounted for by each player on the offensive line. The NFL regulates the fraction of total revenues that can be paid to players. While this is often called a "salary cap", a more accurate description is a ceiling and floor on total payroll. The NFL "salary cap" does not regulate or limit the amount that can be paid to any player; it places a lower and upper limit on the total payroll for players on each team as a fraction of certain revenues earned league wide. Most salaries in the NFL are not guaranteed, and players under contract for multiple years can be released at any time. The "cap value" is the compensation paid to each player that counts toward total team payroll; it includes base salary, a prorated portion of the signing bonus, and incentive bonuses. The "cap value" thus reflects the total compensation of the player in a season, including performance bonuses that depend on the number of plays participated in per season. Cap value can vary depending on the number of plays an individual participates in, so coaches can affect this salary by keeping players on the sideline. Cap value is the standard salary measure in the NFL and has recently been used by Berri and Simmons (2009), Simmons and Berri (2009), Berri et al. (2011) and Keefer (2011). The average salary in the sample is $1.87 million per season. The salary variable exhibits quite a bit of variability and has a long right tail.

In football, there are three offensive line positions: center, guard, and tackle. In our data, some players are identified as generic offensive linemen and not as playing a specific position. Players identified as generic offensive linemen either played multiple

[1] Beginning in the 2002 season, NFL players qualified for performance-based compensation based on playing time. All players qualify for this performance based pay. The largest bonus was $42,048 in 2002 and increased to $299,465 in 2012.

positions on the offensive line, or information about their specific position is not available. The latter case happens in the early years of our data. We create 4 indicator variables to identify linemen playing these 4 positions. Each indicator equals 1 when the player plays that position. From Table 1.1, about one third (35%) of the players are identified as generic offensive linemen in the sample. Notice that creating dummy variables in such way has a limitation. We don't differentiate between left tackle and right tackle here. NFL teams often treat left tackles and right tackles differently. A left tackle is often paid more because he protects the quarterback's "blind side"; since most NFL quarterbacks are right handed, a defensive player approaching the quarterback from his left cannot be easily seen.

Size is an important characteristic for offensive linemen. We collected data on the players height in inches and weight, which ranges from 228 to 375 pounds. The offensive linemen in the sample have a large weight range; the difference between the maximum and minimum weight is 147 pounds. However, weight may not be a good indicator of characteristics which affect performance in this setting.

Like other North American professional sports leagues, the NFL conducts an annual reverse-order entry draft to allocate incoming players to teams. In this sample period, the NFL entry draft consisted of seven rounds. *Overall Draft Selection* is the position where each player was selected in the draft. Players are selected in approximately their order of expected ability. Hendricks et al. (2003) show that draft order is an important measure of the value of a player in the NFL. If a player was not drafted, teams expected that he might not be good enough to make an NFL roster. For undrafted players, we assign 277 to their pick number[2]. We also incorporate into our model an indicator variable *Undrafted*, equal to one of the player was not drafted, together with *Overall Draft Selection*, to capture the effect of not being drafted on salaries, because the difference between the last picked player and undrafted player may not be simply measured by one position in the draft order. 20% of the players in the sample were undrafted.

Better players should start in more games in a season, and receive higher salaries. The variable *Games Started* equals the number of games started by each player in the sample. The average offensive lineman in the sample started ten games per season.

[2]The largest draft number for drafted players in the sample is 276.

Years Experience measures the number of years played in the NFL. For players who entered the NFL through the entry draft, experience is based on the draft year. For players who were not drafted, experience is based on the first year they entered the league.

The variables of interest are the signals of player effort. Stats Incorporated's STATSPASS database contains information about the performance of individual offensive lineman in the NFL. This includes information about the number of penalties committed by each offensive lineman, the number of sacks given up by each offensive lineman, and the yards lost associated with each of these effort signals.

For each player, we divide *penalty yards* by *penalties committed* to estimate *Yards per Penalty*. Average yards per penalty committed reflects both the rate at which penalties are committed and how much the team is punished for the penalties. For example, if two players have the same total penalty yards, say 15, three 5 yard penalties (average 5) may be less harmful to the team than one 15 yard penalty (average 15). The second reason to use average yards per penalty is that reserves play fewer downs than starters, and could have fewer penalties and fewer penalty yards. But a reserve player's average yards per penalty may not be small. The same problem also applies to sacks allowed and sack yards. Thus we also calculate *Yards per Sack* from *sacks allowed* and *sack yards*. Moreover, since offensive line play is highly inter-related, we incorporate other offensive linemen's performance variables into the model. Those variables are *Teammates' Yards per Penalty*, and *Teammates' Yards per Sack* on Table 1.1.

Manski (1993) pointed out a problem that occurs when a researcher observes the distribution of behavior in a population and tries to determine the extent to which this behavior influences the behavior of individuals in this population, the "reflection" problem. We avoid this issue by explaining variation in the salary of an individual offensive lineman using variables that reflect the behavior of all other members of the offensive line excluding that individual. *Teammates' Yards per Penalty* and *Teammates' Yards per Sack* reflect the behavior of other teammates and differs for each individual in the sample.

The average offensive lineman in the sample committed 3.29 penalties per season. In only about 17% of the player-seasons a lineman did not commit a penalty. However, many of those penalty-free player-seasons were by little-used reserves. Among players who started 8 or more games, only about 4% of the player-seasons were penalty free. The maximum number of penalties committed in the sample was 17. The average yards per penalty in the sample was just over 6.

The average offensive lineman in the sample allowed just under three sacks over the course of a season. In about 18% of the player-seasons the player did not give up a sack. Again, this is sensitive to the number of games started. Among players who started at least 8 games in a season, the average number of sacks allowed was four, and only 3% of the player-seasons in the sample were seasons with no sacks given up. The average yards per sack given up in the sample was just over 5.

Table 1.1: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Real Salary (millions)	1.87	1.89	0.089	14.20
Position: Center	0.17	0.38	0	1
Position: Guard	0.21	0.41	0	1
Position: Tackle	0.27	0.45	0	1
Multiple Position Player	0.35	0.47	0	1
Height (inches)	76.4	1.63	72	81
Weight (pounds)	310	20.2	228	375
Years Experience	5.75	3.32	1	20
Overall Draft Selection	137	97	1	277
Undrafted player	0.21	0.41	0	1
Games Started	10.1	6.45	0	16
Yards per Penalty	6.02	3.31	0	15
Yards per Sack	5.28	3.22	0	22
Teammates' Yards per Penalty	7.18	0.77	4.93	10
Teammates' Yards per Sack	6.41	0.94	3	10.37

1.4 Empirical Approach

We analyze longitudinal data on the performance of NFL offensive linemen over ten seasons. We observe individual performance and salary for the same offensive linemen over multiple seasons, so our data constitute an unbalanced panel. Controlling for unobservable heterogeneity, in the form of ability, desire, and other intangible characteristics, represents an important econometric issue in this setting. Researchers typically consider two types of estimators when analyzing panel data with unobservable heterogeneity: fixed effects (FE) and random effects (RE) estimators. Mundlak (1978) points out that the RE estimator assumes all explanatory variables are exogenous and uncorrelated with unobservable individual effects, while the FE estimator allows for all explanatory variables being endogenous and correlated with the unobservable individual effects. Therefore, in the case where some explanatory variables are exogenous and uncorrelated with unobservable individual effects and other explanatory variables are endogenous and correlated with unobservable individual effects, neither the FE nor the RE estimator works well. In this setting, we likely face a situation where some, but not all of the observable variables are correlated with unobservable individual characteristics. Kahn (1993) discusses this problem and the merits of these two estimators in the context of empirical research on earnings in professional sport.

Hausman and Taylor (1981) proposed an estimator appropriate for the case where some of explanatory variables are correlated with unobservable individual effects and others are not. This estimator is based on an instrumental variables (IV) approach. The instruments are constructed using the strictly exogenous variables' between and within variations. This estimator is called the Hausman-Taylor or HT estimator [Baltagi et al. (2003)]. Buraimo et al. (2008) applied the HT estimator to data from professional sport; Dixit and Pal (2010) used it to analyze group incentives.

In general, the estimator developed by Hausman and Taylor (1981) takes the form

$$y_{it} = X_{it}\beta + Z_i\alpha + d_i + u_{it} \tag{1.1}$$

where the subscript i identifies the cross-sectional unit ($i = 1, 2, \ldots, N$), and the subscript t identifies the time periods ($t = 1, 2, \ldots, T$). In this context, the cross-sectional units are players and the time periods are seasons. X_{it} is a vector of time-varying explanatory variables and Z_i is a vector of time invariant explanatory variables. d_i and u_{it} are unobservable random variables that affect the dependent variable y_{it}. d_i is assumed to be $i.i.d.(0, \sigma_d^2)$ and u_{it} is $i.i.d.(0, \sigma_u^2)$. Both are independent of each other and among themselves. d_i is the unobservable individual effect; in this setting, d_i captures a players skill, ability, motivation, or "coachability" that can not be reflected in performance statistics and other unobservable factors that affect the performance of offensive linemen in the NFL, including but not limited to the will to win, morale, leadership, and other intangibles. Both X and Z can be split into two sets of variables, i.e. $X = [X_1, X_2]$ and $Z = [Z_1, Z_2]$. X_1 and Z_1 are exogenous and uncorrelated with both d_i and u_{it}; X_2 and Z_2 are endogenous and correlated with d_i only.

Again, neither the RE nor the FE estimator applies in this setting. The RE estimator, which is basically GLS applied to Equation (1.1), ignoring the endogeneity of X_2 and Z_2, will yield consistent but biased parameter estimates, while the FE estimator, or mean-differencing the explanatory variables in Equation (1.1), eliminates the individual effect d_i as well as the time invariant variables Z_i and hence cannot yield estimates of the vector α, though it can yield consistent estimates of β. The HT estimator resolves this problem.

Hausman and Taylor (1981) develop a variant of the standard IV estimator. It first premultiplies Equation (1.1) by $\Omega^{-1/2}$, where Ω is the covariance matrix of error term $d_i + u_{it}$. After the transformation, it uses a standard two stage least squares IV (2sls) approach with instruments $[Q, X_1, Z_1]$, where Q is for demeaning the variable, specifically $Qy_{it} = y_{it} - \bar{y}_i$. Therefore, the Hausman-Taylor estimator is basically equivalent to perform 2sls using $[\tilde{X}, \bar{X}_1, Z_1]$ as instruments. Intuitively, $\tilde{X}$ can be instrument for X_2 and $\bar{X}_1$ is for Z_2. The advantage of the HT estimator is that all the instruments are derived from within the model. $\tilde{X}$ is the matrix of the deviation of X (both X_1 and X_2) from its associated mean. $\bar{X}_1$ is the mean of the exogenous time-varying variables. These are the standard within and between components of the FE and RE estimators.

Another important issue associated with the HT estimator is the identification condition for the model. As pointed out by Baltagi et al. (2003), if the number of exogenous time-varying variables X_1 is greater or equal to the endogenous time-invariant variables Z_2, the model is identified and the HT estimator is more efficient than the FE estimator. If the condition fails, the model is under identified and the HT estimator cannot outperform FE and the coefficient on Z, α, cannot be estimated.

The specific form of the HT estimator in this case is:

$$\ln salary_{it} = X_{1it}\beta_1 + X_{2it}\beta_2 + Z_{1i}\alpha_1 + Z_{2i}\alpha_2 + d_i + u_{it} \tag{1.2}$$

where X_{1it} is a vector of $n \times k_1$ exogenous, time-varying explanatory variables that are assumed to be uncorrelated with d_i; X_{2it} is a vector of $n \times k_2$ endogenous, time-varying variables that are assumed to be correlated with d_i; Z_{1i} is a vector of $n \times l_1$ exogenous, time-invariant explanatory variables that are not correlated with d_i; Z_{2i} is a vector of $n \times l_2$ endogenous, time-invariant variables that are correlated with d_i; and d_i is a player-specific effect. The model is identified when $k_1 \geq l_2$.

Note that X_{1it} contains a vector of variables identifying the team that each lineman played for over the sample period. This variable captures unobservable heterogeneity across teams. It is time varying because some players change teams in the sample period. Unobservable team specific heterogeneity may arise from various sources, such as culture, managerial style, owner and coach preferences and factors related to the city where team plays. The general business environment, stadium deal, and residents' enthusiasm for football in the city could all be captured by this team effect.

Again, we have unbalanced panel data; the cross-sectional unit is an individual player. The dependent variable is the natural logarithm of the real salary in each season. The model contains 43 variables classified in four categories (exogenous time-varying, endogenous time-varying, exogenous time-invariant and endogenous time-invariant). The vector of exogenous time-varying variables contains 34 variables, including 32 team dummy variables, *Years Experience* and its squared term, *Years Experience*^{2}, which together measure the effect of experience. All these variables are assumed to be independent of the unobservable individual effects. It is unlikely that unobservable individual effects (for example, ability, will to win or drive) influence the player's team or experience. Though a talented player is expected to play longer in the league, many talented players retire early because of injury. The typical career is short, normally less than 10 years. Some relatively less talented players can stay in the league for a relatively long time as long as they can stay healthy. We add both *Years Experience* and *Years Experience*^{2} into the model following the same logic of the quadratic relationship between salary and age, and hence we expect a positive coefficient on *Years Experience* and negative coefficient on the squared term.

The vector of endogenous time-varying variables includes 6 variables: an indicator variable identifying linemen who played multiple positions over the course of the season, the number of games started in the season, and four variables reflecting the effort signals received by the team: yards per penalty, yards per sack allowed, teammates yards per penalty, and teammates yards per sack allowed. Linemen who played multiple positions are likely to be reserves filling in occasionally at many positions, perhaps because they lack the ability to start at a position full time. We expect the coefficient on this variable to be negative. Better players should start in more games in a season, so *Games Started* should have a positive coefficient.

Again, signals received about individual players (yards per penalty and yards per sack) reflect both effort decisions and a common shock experienced by all offensive linemen on a team. Penalty yards and sack yards reflect shirking by a player, so those variables are expected to have negative coefficients, based on standard efficiency wage theory. The signals received about other players effort (teammates' yards per penalty and teammates yards per sack) are motivated by the model of repeated interaction among workers on a team developed by Che and Yoo (2001) and discussed above. The estimated parameters on these variables will indicate the presence of JPE wage schemes, if negative, and RPE wage schemes, if positive.

The performance variables are all likely to be correlated with the unobservable individual effect. For example, if a player works very hard in both practice and games, he will probably perform well in games and the good performance would in turn affect his desire to win and/or other unobservable factors.

The only exogenous time-invariant variable is *Height*, which is constant over time and can not be affected by unobservable individual effects. Conversely, it is also unlikely that height would influence ability or other unobservable individual effects. Height may be a basic physical qualification to become an offensive lineman. But conditional on becoming a linemen, we can't say taller players are superior to shorter players and the players themselves would not believe that being a few inches taller would bring them some advantage over other players. The unobservable individual heterogeneity should be uncorrelated with height. However, among offensive linemen, generally speaking, tackles are tallest, then guards, and centers are often the shortest players. Also the teams usually pay tackles, especially left tackles, more. So it seems taller players earn more money and hence we tentatively expect a positive coefficient on *Height*. Note that in our data, about one quarter of players switch positions over their careers. So height and position are not highly correlated.

The two endogenous time-invariant variables are *Undrafted* and *Overall Draft Selection*. These two variables, especially the second one, are regarded as important indicators of players' ability and hence are expected to have significant effects on salary. Moreover, since draft order reflects individual ability, this variable is likely to be correlated with the unobservable individual effects. A player who is believed to have greater ability and potential will be selected earlier in the draft. Those factors will in turn influence attitude toward the game and effort. Since a smaller draft number means a better player, we expect *Overall Draft Selection* will have a negative coefficient. But for *Undrafted*, the sign could be either negative or positive, because the effect of being undrafted is captured by both these variables. If the sign is positive for *Undrafted*, the coefficient on *Overall Draft Selection* plus the coefficient on *Undrafted* should be negative, in which case undrafted players earn less than the last drafted player.

In our model, $k_1 = 34$ and $l_2 = 2$, even after omitting one team indicator variable, k_1 is still much larger than l_2. Therefore, our model is identified.

1.5 Results and Discussion

The parameter estimates, standard errors and P-values obtained from the HT estimator applied to Equation (1.2) are shown on Table 1.2. Note that we do not report the parameter estimates from the 31 team indicator variables since their effects are not our primary interest here. These results are available by request from the authors.

The relationship between experience and salary takes the standard hump-shaped form, first increasing as human capital accumulates and then decreasing as it depreciates. Declining physical ability, and the cumulative effect of injury and physical wear-and-tear also contribute to the decline in earnings as experience increases in this setting. As expected, linemen who play multiple positions earn less than those who play a single position, even holding games started constant. This effect may also reflect returns to specialization among offensive linemen in the NFL. Starting linemen

Table 1.2: Hausman-Taylor Regression Results, NFL Offensive Linemen 2000-2009

Dependent Variable: Log(Real Salary)	Coefficient	Std. Err.	p-value
Variable Type: Time Varying Exogenous			
Years Experience	0.481	0.015	$<$0.001
Years Experience^2	-0.024	0.001	$<$0.001
Variable Type: Time Varying Endogenous			
Multiple Position Player	-0.158	0.046	0.001
Games Started	0.017	0.003	$<$0.001
Yards per Penalty	-0.010	0.004	0.014
Yards per Sack	-0.001	0.004	0.841
Teammates' Yards per Penalty	0.015	0.015	0.333
Teammates' Yards per Sack	-0.025	0.013	0.045
Variable Type: Time Invariant Exogenous			
Height	0.154	0.059	0.009
Variable Type: Time Invariant Endogenous			
Overall Draft Selection	-0.004	0.002	0.007
Undrafted	-0.017	0.504	0.972
Observations	2297		
Individuals	611		

earn a significant premium over reserves who start no games, or perhaps only a few games, per season.

The variables of interest are the effort signals for the two effort signals, committing penalties and allowing sacks. Again, the salary variable reflects a signing bonus, base salary, and performance related pay that varies systematically with the number of plays participated in. Effort signals can affect salary if coaches make substitutions based on these signals over the course of a game, generating differences in the number of plays each player participates in over the course of a season.

The results exhibit an asymmetric pattern that supports the idea that a JPE wage scheme exists in this setting, as one effort signal from teammates, giving up sacks, affects the salary of other players. Each additional yard per sack given up by teammates reduces the earnings of an individual NFL offensive lineman by about 2.5%. However, yards per sack given up by an offensive lineman do not reduce his own earnings. This pattern of results suggests that a low powered mechanism exists through which offensive linemen can punish teammates for shirking. In the repeated interaction that takes place on an NFL offensive line over a season, if teammates observe another lineman shirking, they can punish him by giving up sacks without reducing their own salary, providing a low power mechanism for enforcing the a subgame perfect equilibrium in which these workers supply effort in each period, just as the model developed by Che and Yoo (2001) predicts. Note that sacks have less impact than offensive holding, a penalty that might be committed to avoid a sack; offensive holding results in a 10 yard penalty while the average sack in the 2002-2010 NFL seasons resulted in a 6.37 yard loss (median 7 yard loss). This result is consistent with the use of JPE wage schemes based on sacks allowed among teams of workers on the offensive line.

The other effort signal, penalties committed by players, does not appear to be a viable mechanism to punish other workers who shirk. The yards per penalty committed by other teammates does not affect the earnings of individual linemen in this sample. However, the larger the yards per penalty committed by an NFL offensive lineman, the lower that lineman's salary, other things equal. Recall that, unlike sacks allowed, the commission of penalties in the NFL is monitored by an independent party, the team of referees that officiate NFL games. Each additional yard of penalties committed reduces the earnings of the offensive lineman by 1%. This result can be motivated by the model developed by Kvaløy and Olsen (2012), which includes indispensable human capital in teams. Alternatively, standard efficiency wage models, or principal-agent models with monitoring by the firm, in which firms penalize workers who shirk with lower wages if they can observe this shirking, can also motivate this result.

The results indicate that a salary premium to height exists in this setting. Taller offensive linemen earn more than their shorter teammates, other factors constant. The premium to height is substantial, about 15% per inch. The premium could be due to higher productivity among taller linemen. Persico et al. (2004) find a wage premium to height in the general population and attribute this to events in early adulthood. Schultz (2002) finds a wage premium to height and attributes it to human capital accumulation. These two explanations could also hold in this population.

The results also show that linemen drafted higher in the entry draft earn a higher salary, even years after being drafted and controlling for experience. This persistence of draft position in salaries was also documented by Hendricks et al. (2003) in the NFL. Note that 65% of the variation in the dependent variable can be attributed to unobservable player-specific heterogeneity captured by the random effects term in Equation (1.1).

1.5.1 Robustness Checks

Table 1.2 contains results for a single model specification. The HT estimator can be sensitive to the specification of the time-varying endogenous and exogenous variables. No reliable, commonly used tests exist to provide guidance about which variables belong in which category in the HT estimator. We performed a number of robustness checks on the results reported on Table 1.2. The results, specifically the sign and significance of the effort signal variables, were not sensitive to these changes to the model.

We added a series of indicator variables for each season in the sample, to capture any unobservable heterogeneity in offensive line play that would affect compensation, sacks allowed and penalties committed. These factors could include changes to the specific rules about what offensive linemen can and cannot do, and any changes in the enforcement of the existing rules that might vary systematically across seasons. The inclusion of this vector of indicators to the time varying exogenous variable list had no effect on the parameter estimates of interest. We also added an indicator variable for changes in the head coach of the team; this also had no effect on the results.

Equation (1.1) does not control for systematic variation in the salaries of offensive linemen across positions on the line. If sorting by ability takes place, say the most talented linemen become tackles, or the smartest linemen become centers, then this needs to be controlled for in the regression model. We included a vector of variables identifying the specific position played by the linemen in our sample, both as an time varying exogenous and time varying endogenous variable. The inclusion of these indicator variables had no effect on the results. Similarly, changing the indicator variable for a lineman who played multiple positions from time varying endogenous to time varying exogenous had no effect on the results.

We also added BMI (body mass index) to the model as a time-varying endogenous variable. Although height does not change, weight does. BMI could affect agility, or other factors that affect the play of a lineman. Adding BMI, and BMI^2 had no effect on the results.

The relationship between sacks allowed and the salary of offensive linemen could be affected by some confounding factors like the tendency of the team's offense to pass, the overall efficiency of the team's offense, or some other characteristic of the offense that affects the team's overall success. The results reported on Table 1.2 were robust to the inclusion of a variety of variables like the fraction of plays that were passes, yards per pass, the fraction of passes completed, and first down efficiency to the list of time varying endogenous variables.

Differences in team financial conditions might systematically affect the effect of the effort signal on player salaries. For example, teams with lower revenues might punish players with many negative shirking signals more than teams with higher revenues. The results reported on Table 1.2 were robust to the inclusion of a variety of variables reflecting the revenues earned by teams, and the fraction of the team's revenues paid to players.

The group size of offensive linemen in a team may affect the interactions among these linemen. The repeated interactions could be different for a small group of offensive linemen and a large group of offensive linemen. The results reported on

Table 1.2 were robust to the inclusion of the number of offensive linemen in the team.

We also estimated a fixed effect model version of Equation (1.1); the results were similar to the Hausman-Taylor results on Table 1.2 for the effort signal variables. Note that the 3 time-invariant variables *Height*, *Overall Draft Selection* and *Undrafted* were omitted due to collinearity. As mentioned in empirical approach section, the FE estimator can yield consistent estimates of time varying variables by eliminating the individual effect and time-invariant variables.

1.5.2 Evidence From Play-by-play Data

The regression results in the previous section suggest that giving up a quarterback sack represents a low powered effort signal in a joint performance evaluation wage scheme. In order for giving up a quarterback sack, a form of shirking, to be a low powered signal, it must be possible for NFL offensive linemen to send such signals at points in the game where the game outcome will not be too adversely affected. For example, sending such a signal when the player's team is about to score a tying or go ahead touchdown would not be low powered, since it could lead to the team losing a winnable game. Also, not all sacks allowed are shirking-related signals; in some cases, sacks are unavoidable consequences of the game situation and personnel.

In order to assess the viability of allowing sacks as a shirking-related signal in a JPE wage scheme, we analyzed play level outcome data from all regular season games in the 2002 through 2010 seasons.[3] The sample contains information about 305,483 individual plays. We removed kickoffs, field goal attempts, punts, extra point attempts and two-point conversion attempts, as these plays could not result in a quarterback sack. We identified plays resulting in a quarterback sack, and the number of yards lost, from the remaining plays. A sack was defined as a play in which the quarterback was tackled at or behind the line of scrimmage. Plays where the quarterback advanced the ball for positive yardage were not identified as sacks. 10,394 of these 305,484 plays resulted in a quarterback sack for an average frequency of 3.4 sacks per 100 plays. The average sack resulted in a loss of 6.37 yards; the

[3]These data come from `http://www.advancedfootballanalytics.com/2010/04/play-by-play-data.html`.

longest sack in the sample resulted in a loss of 38 yards.[4] Unfortunately, we cannot link specific offensive linemen to a sack in this data set. Only the season-level data set analyzed in the previous section links specific offensive linemen to quarterback sacks allowed. This data source does contain information on the exact game conditions in terms of time, field position, down and distance, and score, for each play during the NFL season.

We identify games and game situations where giving up a quarterback sack would not have a large impact on the game outcome, or season outcome for teams. In these situations, giving up a quarterback sack would have a low cost to the team in terms of the effect of this signal on the game or season outcome, a low powered signal.

We identified five different situations in which a quarterback sack might not affect game or season outcomes: games early in the season, games with large point spreads where one team is much stronger than the other team, plays during games where one team has a large lead, plays late in games where one team has a large lead, and games where one of the teams has been eliminated from playoff contention. Sacks during these games, or plays during games with these conditions, could be considered low powered signals.

Table 1.3 summarizes the frequency and losses for quarterback sacks in each of these low and high power game situations. The first two columns are for high power game situations and the second two columns are for low power game situations. The first case is games early and later in the season. We divided games into the first 4 games of the season, and the last 12 games of the season. Games early in the season may be less important than games later in the season. From the top panel of Table 1.3, 228,480 plays took place in games 5 through 16 in this sample, and 77,003 plays took place in the first four games. The frequency of sacks allowed was identical in these two periods, 3.4 sacks per 100 plays. The average loss on each sack was similar. There were 7,756 sacks in games 5-16 and 2,638 sacks in games 1-4 in these seasons.

The second low power setting is plays when one team has a large lead on the other team. We define a large lead as a lead of 14 points or more. From the second panel on Table 1.3, of the 305,483 plays analyzed, 60,875, just under 20% of the

[4]Philadelphia quarterback Donovan McNabb versus the Oakland Raiders on 18 October 2009 in the second quarter when trailing by 7 points.

Table 1.3: Sack Statistics for Low and High Power Game Situations

	Mean	Number	Mean	Number
	Games 5-16		Games 1-4	
Sack frequency	0.034	228480	0.034	77003
Yards lost per sack	-6.39	7756	-6.34	2638
	Score difference ≤13		Score difference ≥14	
Sack frequency	0.033	244608	0.037	60875
Yards lost per sack	-6.34	8170	-6.52	2224
	Score difference ≤13, 4th Q		Score difference ≥14, 4th Q	
Sack frequency	0.034	277734	0.037	27749
Yards lost per sack	-6.34	9359	-6.71	1035
	Point spread ≤9		Point spread ≥10	
Sack frequency	0.034	259200	0.034	46283
Yards lost per sack	-6.37	8798	-6.42	1596
	Teams with ≤9 losses		Teams with ≥10 losses	
Sack frequency	0.034	268810	0.034	36673
Yards lost per sack	-6.38	9132	-6.35	1262

plays, took place when one team had a 14 point lead or larger. The sack frequency was slightly higher when one team had a large lead, and the average sack resulted in a slightly larger loss, 6.52 yards, compared to the average loss during closer game situations. A substantial number of sacks took place during these low powered game situations, providing offensive linemen with ample opportunities to send low powered effort signals. The third panel summarized sacks in the fourth quarter when one team has a large lead. It should be an even more low powered situation, since little time remains in the game for the losing team to stage a comeback. The 27,749 plays run in this condition represents 9% of the plays in the sample. Again, the sack frequency is higher and the average sack results in a larger loss in this relatively low powered situation.

The fourth low powered setting analyzed are games where one team is much stronger than the other team. Since the weaker team is less likely to win a game, games with a large point spread could represent a low powered setting since the stronger team is much more likely to win the game no matter how many sacks are given up. We obtained data on the closing point spread for all NFL regular season games over the 2002-2010 seasons. We defined games where one team is perceived as much stronger than the other as games when one team was favored by 10 or more points. This occurred in about 15% of the games played in this period. The rate of sack frequency was the same in games with a heavy favorite and games with no favored team or games with a smaller point spread. The average loss per sack was similar in the two groups of games.

Finally, games involving a team with no chance of making the playoffs could be low powered settings, since one of the teams has a reduced incentive to win the game. During this period, no team with more than 9 losses qualified for the NFL post season, so we split the sample into game involving at least one team with 10 ore more losses and games with no teams with 10 or more losses. Games with one or more teams with 10 or more losses take place late in the 16 game NFL season. From the last Panel on Table 1.3, the sack frequency and average loss per sack were very similar in these two groups of games. 1,262 of the 10,394 sacks in the sample took place in games where one or more teams had 10 or more losses.

From Table 1.3, NFL offensive linemen have substantial opportunities to send low powered effort signals by allowing a quarterback sack, over the course of an NFL season. Thousands of quarterback sacks take place in games, or game settings, where a quarterback sack is unlikely to have a large impact on the game or season outcome for the team.

1.6 Conclusions

Che and Yoo (2001) and Ishida (2006) developed models that describe the optimal incentives in a setting where a team of workers repeatedly interact in producing a specific type of output. These models conclude that when workers on a team interact repeatedly, firms have an incentive to use JPE wage schemes, as they result in lower incentive costs when workers can monitor the effort decisions made by other team members and use a low powered punishment mechanism to deter other team members from shirking.

Offensive line play in the NFL contains a number of institutional characteristics that closely resemble the setting for the models developed by Che and Yoo (2001) and Ishida (2006). These workers interact repeatedly, can monitor each other, often motivate each other, and have significant autonomy to carry out their tasks during the course of a game; in addition, relative performance evaluation occurs in this setting. We collected data on two plausible effort signals that teams receive from offensive linemen: penalties committed, which can reflect shirking and are monitored by a group of independent agents, the officials in NFL games, and sacks allowed, which are monitored only by the players and the coaches of the team. We find that sack yards allowed by other linemen on the team significantly reduce the earnings of offensive linemen, but sack yards given up by each lineman does not reduce his own earnings, other things equal. These results are consistent with the presence of a low power punishment mechanism through which a group of team members can punish another team member who shirks during their repeated workplace interaction. The results are also consistent with the presence of JPE wage schemes in this setting. Both are features of the models developed by Che and Yoo (2001) and Ishida (2006), so the results here suggest that these models explain observed outcomes in the performance and earnings of NFL offensive linemen.

These results also extend the growing literature of the importance of peer effects in sport. Depken (2000) found evidence supporting peer effects in Major League Baseball. Guryan et al. (2009) found no evidence of peer effects in professional golf, based on a setting that includes random assignment of team mates. Depken and Haglund (2011) found evidence of peer effects in foot races. This analysis uses a novel setting, offensive line play in the NFL, which exploits the repeated nature of productive interaction in teams, and which generates plausible observable effort signals. Both these features are unique to the literature on peer effects in teams.

Finally, the presence of JPE wage schemes and viable weak powered enforcement mechanisms have important implications for research on the payroll-success relationship in professional sports. Many models in this literature assume that higher payrolls lead to increased production of wins; see, for example, the model developed by Fort and Quirk (1995). Che and Yoo (2001) show that JPE wage schemes generate incentives at a lower cost than RPE incentive schemes. If JPE wage incentive schemes are widespread in professional sports, then the relationship between total payroll and team success may not be as well-behaved or strong as many existing models of team production in sport assume, especially if some teams use JPE wage schemes and others use RPE schemes.

Chapter 2

Consumption Commitments and Simultaneous Insuring and Gambling: Evidence from Canada

2.1 Introduction

For more than half a century, economists have studied a seemingly contradictory economic decision: the simultaneous purchase of insurance and gambling. This behavior is considered contradictory because, according to the standard von Neumann-Morgenstern expected utility model, individuals should either purchase insurance or gamble, but should not engage in both activities simultaneously. According to expected utility theory, those who purchase insurance are risk averse with a concave utility function. Those who gamble, on the other hand, are risk seekers with a convex utility function.

In pioneering work, Friedman and Savage (1948) proposed that local non-concavity of the utility function provides one explanation for this contradictory behavior. Their paper inspired an extensive literature, including Dowell (1985), Dowell and McLaren (1986), Chetty and Szeidl (2007) and Jones (2008). This paper extends this line of research theoretically and empirically. It explores the theoretical implications of "consumption commitments" proposed by Chetty and Szeidl (2007) and empirically tests their theory using Canadian household survey data.

According to the model developed by Friedman and Savage (1948), an individual's utility function has both convex and concave regions. When the individual's income is around the transitional region from the concave part to the convex part, she might purchase insurance and gamble at the same time. In this particular region of the utility function, the individual's decision to gamble is based on the convex section and the decision to purchase insurance is based on the concave section. This theory has not escaped criticism. For example, Markowitz (1952) argues that Friedman and Savage's model predicts that individuals with very high income will not gamble at all because the final part of the utility function must be concave, but evidence shows that people purchase insurance and gamble at all wealth levels. Also, a non-concave utility function violates the assumption of diminishing marginal utility.

Later research (Kwang, 1965; Jones, 2008) posits that local non-concavity can arise from the presence of expensive indivisible consumption goods such as homes, cars or vacations. Indivisibility in consumption set can lead to a utility function that is consistent with Friedman and Savage's model and does not violate diminishing marginal utility. More specifically, a local non-concavity can be caused by discrete changes in utility generated by the consumption of indivisible goods. Dowell (1985) and Dowell and McLaren (1986), however, argued that many indivisible goods can be divided through leasing, installment payments, or other types of borrowing and lending arrangements. Alternatively, they explain local non-concavity using the effect of wages and work-leisure choice on lifetime earnings and show that, at a critical wealth level, a change in labor supply from a maximum to zero (for example, retiring from work) yields a discrete increase in utility, due to the complementarity between goods and leisure.

Chetty and Szeidl (2007) offer an alternative explanation. They suggest that "consumption commitments" play an important role in consumer's behavior under uncertainty. Consider an individual consuming two kinds of goods: a freely adjustable good and a good that is costly to adjust because of some commitment. Housing represents one commitment good. Most individuals must pay either a mortgage or rent every month to acquire housing services. Chetty and Szeidl (2007) show that commitments amplify risk aversion over moderate uncertain payoffs, which in turn changes the shape of the concave value function over wealth, generating Friedman-Savage style local non-concavities. Unlike Friedman and Savage (1948), since commitments are endogenous to wealth, local non-concavities in this model can occur anywhere along the value function. Therefore incentives to gamble can occur at different wealth levels. Furthermore, this approach does not violate the assumption of diminishing marginal utility.

Chetty and Szeidl (2007) focus on moderate gambles and the impact of commitment on risk aversion. Here, this paper extends their model by taking into account activities with large uncertain payoffs like lotteries and by examining the impact of commitment on risk loving. According to Chetty and Szeidl (2007), the size of a gamble increases in the standard deviation of its possible outcomes. Lotteries, which often pay jackpots of millions of dollars, are good examples of activities with large uncertain payoffs. Casino gambling, slot machines and video lottery terminals that normally offer smaller prizes than lotteries, can be considered as moderate-to-large uncertain payoffs, and are referred to as a *small prize gamble.* Insurance, which is purchased to reduce risk, has a fairly small standard deviation and is regarded as an activity with a moderate uncertain payoff.

This paper first explores the theoretical implications of commitments and risk loving over different uncertain payoffs. Commitments magnify risk loving over large uncertain payoffs. For example, people with more commitments are more likely to play the lottery, a form of gambling with relatively large uncertain payoffs. For moderate-to-large uncertain payoffs, the impact of commitments exhibit a twofold pattern, depending on the adjustment costs associated with the commitments. Specifically, moderate commitments, like insurance premium payments or vehicle financing installment payments, magnify risk loving over this type uncertain payoffs; but large commitments such as large mortgage payments mitigate risk loving.

The second contribution of this paper is empirical. Much of the existing literature on simultaneous insuring and gambling is theoretical. Far less attention is given to empirical analysis of individuals' participation in gambling and insurance purchases. This paper addresses this gap in the empirical literature by analyzing confidential micro-level data from the Canadian Survey of Household Spending (SHS) over the period 2004-2009 using a multivariate probit model to simultaneously estimate individual participation in government-run lotteries, small prize gambles like casino gambling, slot machines, or video lottery terminals (VLTs), and life insurance. Government-run lottery is interpreted as gambling with a large uncertain payoff and small prize gambling a moderate-to-large uncertain payoff. Life insurance is interpreted as a moderate payoff activity. Controlling for common observable and unobservable determinants of gambling and insurance purchase, we find evidence consistent with the theoretical predictions of the model. Specifically, households who have greater commitments are more likely to buy lottery tickets and life insurance. For small prize gambles, the impact of commitments is size-dependent, as predicted by the theory.

2.2 Relevant Literature

There are several other models, besides the models with local non-concavity in utility functions discussed above, which offer potential explanations for simultaneous insuring and gambling. Examples include market imperfections (Kim, 1973), risk allocation over time (Eden, 1977), prospect theory (Kahneman and Tversky, 1979), moral hazard (Szpiro, 1992), the utility of gambling (Conlisk, 1993), and the utility from time saved (Nyman et al., 2008).

Kim (1973) developed a model of decision making under uncertainty that does not rely on assumptions about the curvature of utility function. Kim (1973) posits that an individual maximizes expected total gains from gambling or minimizes expected total losses from buying insurance over his lifetime. The individual evaluates the gains (or losses) using interest rates over the remaining life span. Likewise, the costs of gambling or buying insurance are calculated with the inclusion of forgone interest income that the individual would have earned over his lifetime if he does not use it to pay for gambling or insurance. Under certain market imperfections, the interest rates used to discount the gains or losses and those used in the cost calculation can differ from each other. Kim argues the simultaneous purchase of insurance and gambling would happen when interest rate for the expected gain (or loss) is greater than that for the cost. This explanation, however, can lead to an unrealistic prediction that all the people should both buy insurance and gamble under market imperfections.

Eden (1977) analyzed the role of insurance and gambling in allocating risk over time. He showed that society as a whole will eliminate future uncertainty regardless of individuals' attitudes toward risk. Risk seekers will engage in gambling to create the present risk after eliminating future risk by purchasing insurance. Further, Eden (1980) argues that the outcome of a gamble is usually known in a short time while insurance is purchased to mitigate the consequences of adverse events that occur in the future. The risk preference for present gambles is not necessarily the same as that for future gambles. Risk seekers who engage in gambling could also have positive demand for information about wealth realization in the future and would, therefore, also purchase insurance.

Szpiro (1992) employs a hyperbolic absolute risk aversion (HARA) utility function to model insurance purchasing as an expected utility maximization problem. He shows that while risk averse individuals buy full insurance, partial insurance and even sell insurance depending on uninsured amounts and values of a loading factor, risk seekers will over-insure. In other words, risk seekers are willing to pay an additional premium above the premium demanded to insure a whole asset. The additional premium payment, which entitles them to an extra payoff, is actually equivalent to the purchase of a lottery ticket which pays off upon a loss producing event. A risk seeker therefore tries to exploit this form of moral hazard. Szpiro (1992) provides a simpler explanation for the purchase of insurance by a risk seeker. However, the model cannot explain why gamblers do not pay an additional premium when they buy insurance.

Another important theory, prospect theory, developed in Kahneman and Tversky (1979), stands out as an alternative to models based on expected utility. According to prospect theory, people make decisions under uncertainty by evaluating potential gains and losses instead of real probable outcomes. Outcomes above a certain reference point (often the current asset position) are gains while outcomes below the reference point are losses. Further, people are more sensitive to potential losses than gains, which is called loss aversion. The value function of outcomes takes an S-shape and is asymmetric. Prospect theory holds that people tend to overweight small probability events and underweight events with moderate to high probabilities. This theory generates a "fourfold pattern of risk attitudes": when the event probabilities are moderate to high, people are risk averse for gains but risk seeking for losses; when the events probabilities are low, people behave as risk seekers for gains but are risk averse over losses. A well-known problem of prospect theory is the violation of stochastic dominance [Wakker (2010), 153-154.]. To resolve this problem Tversky and Kahneman (1992) develops *cumulative prospect theory*, which applies weight to the cumulative probability distribution function rather than the probabilities of outcomes. This leads to one major difference from original prospect theory that people tend to overweight extreme events with small probability, rather than to overweight all small probability events regardless of the outcomes. More importantly, cumulative prospect theory can also handle ambiguity with unknown probabilities.

Many criticisms of prospect theory exist. Conlisk (1993) notes that prospect theory is restricted to small probability events. It does not fully account for the simultaneous insuring and gambling involving moderate probability events. Conlisk proposes a theory that decision making under risk is not only based on the expected monetary utility but also on the utility of gambling itself. He appends an extra term to the conventional expected utility function to capture the utility of gambling. This utility of gambling can be interpreted as the utility of excitement and entertainment from gambling. Because of this extra utility term, a risk averse individual would also gamble. The utility-of-gambling theory, however, is criticized in Diecidue et al. (2004) for the implied violation of transitivity or stochastic dominance.

Finally, Nyman et al. (2008) suggest that people purchase lottery tickets to obtain "something for nothing". Because of scarce resources, an individual usually has to give up something to obtain something else. But in the case of lottery ticket purchase, gamblers are trying to obtain some additional income for almost nothing. Nyman et al. (2008) further argue that utility from playing the lottery is generated not only from the money utility of the prize, but also from the utility cost saved by not having to work longer if the prize is won. This additional utility resembles a counterpart to the Conlisk (1993) utility of gambling model but has a different interpretation. Since total utility includes the utility cost saved by not needing to work longer, work experience matters when an individual decides to gamble. Individuals with a full-time job are more likely to play the lottery than part-time workers and a part-time worker is more likely to play the lottery than the unemployed.

None of these papers examine the effect of consumption of commitment goods on gambling. Recent research has emphasized the importance of commitment goods in decisions under uncertainty. The next section examines the role played by commitment goods in the decision to gamble.

2.3 A Model of Commitment Under Uncertainty

2.3.1 Model Setup

Following Chetty and Szeidl (2007), assume a household only consumes two goods. One is freely adjustable at all times, and the other is costly to adjust. The freely adjustable good (g) can be food, and the costly-to-adjust good (z) can be housing, vehicles, or insurance. The consumption of this type of good requires households to make a commitment, for example, paying the mortgage or insurance premium regularly and on time. Therefore, we call them commitment goods. Households have to pay an adjustment cost $k \cdot z$ ($k > 0$) to change their consumption of a commitment good, for example, moving to a new house or cancelling an insurance policy. Assume the household lives for T periods and maximizes expected lifetime utility:

$$E_0 \Sigma_{t=1}^{T} u(g_t, z_t) \tag{2.1}$$

subject to a dynamic budget constraint:

$$W_{t+1} = W_t + y_{t+1} - g_t - z_t - kz_{t-1} \cdot 1\{z_{t-1} \neq z_t\}, \tag{2.2}$$

where y_t denotes (risky) income in period t and W_t is household wealth at the beginning of period t. A terminal condition $W_{T+1} = 0$ applies.

Chetty and Szeidl (2007) assume a zero discount factor and interest rate for simplicity; under these conditions the optimal consumption path for a household is flat: $g_t = g_1$ and $z_t = z_1$ for all t. They further argue that households who abandon commitments would only do this in the first period, otherwise they would never abandon their commitments. Then the household's problem becomes maximizing the value function in period 1, $v^c(W, z_0)$, which is

$$v^c(W, z_0) = \max(v^0(W, z_0), v^m(W, z_0)), \tag{2.3}$$

where $v^0(W, z_0)$ is the utility generated by not abandoning commitments and $v^m(W, z_0)$ is the utility generated by abandoning commitments in period 1.

Chetty and Szeidl (2007) show that a household will not change the consumption of a commitment good for all $W \in (s, S)$ and $s < S$, and will change when $W \notin (s, S)$. The decision largely depends on the magnitude of the wealth shocks, or uncertain payoffs, households experience. Suppose a household receives a large positive wealth shock, for instance, winning a large lottery jackpot. If the household does not want to move to a larger house, for example, an alternative would be to increase consumption of food significantly, which generates a limited utility increase because the marginal utility quickly diminishes. This represents a worse scenario relative to moving and paying the adjustment cost. In contrast, if the household experiences a small wealth shock, it is preferable to maintain the commitment. The utility gain from changing consumption cannot compensate for the loss from the adjustment cost. Chetty and Szeidl (2007) further show that the (s, S) band expands when the adjustment cost increases, i.e. if $k_2 > k_1$, then $S^{k_2} > S^{k_1}$ and $s^{k_2} < s^{k_1}$.

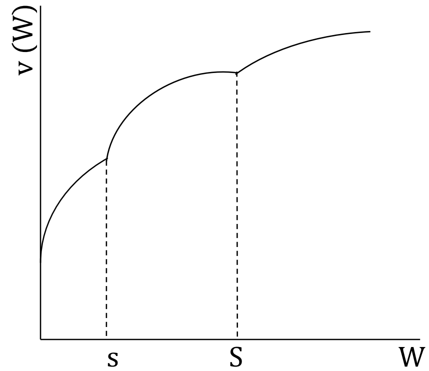

Figure 2.1: Value Function

To further explore household risk preferences, denote the value function of a household without commitment by $v^n(W)$. For the no commitment case, the constant relative risk aversion (CRRA) parameter is $\gamma^n(W) = -v^n_{WW}W/v^n_W$ and the CRRA parameter for a household with commitment is $\gamma^c(W)$.

Chetty and Szeidl (2007) show that the presence of commitment goods magnifies household risk aversion at certain wealth levels: $\gamma^c(W) > \gamma^n(W)$ for all $W \in (s, S)$, and risk aversion is higher for all wealth levels $W \in (s, S)$ than $W' \notin (s, S)$, namely, $\gamma^c(W) > \gamma^c(W')$.[1] Figure 2.1 illustrates the shape of the value function $v^c(W, z_0)$. There are locally non-concave regions, similar to the Friedman and Savage (1948) utility function, in the neighborhood of $W = s$ and $W = S$. Therefore, households at certain wealth levels have an incentive to gamble. The expected utility of gambling is higher due to the convex curvature of the value function. Intuitively, households gamble for an opportunity to drop their prior commitments and to enter into new commitments. For example, they would like to participate in gambling activities that may help them upgrade to a larger house or buy a nicer car. They are not gambling to purchase additional burger or sandwich.

It is of more interest to analyze the effect of commitments with different adjustment costs on risk aversion. Following Chetty and Szeidl (2007), a random variable $\tilde{W}$ with finite expected value $E\tilde{W}$ is used to represent a realization of lifetime wealth.

[1] Chetty and Szeidl (2007) argue this result needs another condition, $W' > kS/T$, because the reduction made by adjustment costs at very small wealth levels can also increase risk aversion. Therefore, there is a lower bound for the wealth level.

We interpret $\tilde{W}$ as reflecting the household wealth level after gambling. It is the result set of facing an uncertain payoff or shock. Let $W^{CE}(\tilde{W}, z_0, k)$ represent the certainty equivalent for a household with adjustment cost $k \cdot z_0$. The household is indifferent between a gamble that generates $\tilde{W}$ and wealth generated by a secure income of $W^{CE}(\tilde{W}, z_0, k)$. The proportional risk premium is hence defined by $\pi(\tilde{W}, z_0, k) = 1 - W^{CE}(\tilde{W}, z_0, k)/E\tilde{W}$, and reflects the welfare cost of the uncertain future payoff. Chetty and Szeidl (2007) introduce the concept of "equivalent relative risk aversion," denoted by $\gamma(\tilde{W}, z_0, k)$, to represent the CRRA parameter for a household with certainty equivalent $W^{CE}(\tilde{W}, z_0, k)$. They interpret equivalent relative risk aversion as "an approximate measure of the risk premium per unit" of uncertain payoff. It is defined as:

$$\gamma(\tilde{W}, z_0, k) \approx 2\frac{-\log(1 - \bar{\pi}(\tilde{W}, z_0, k))}{\text{var}[\log \tilde{W}]}. \tag{2.4}$$

Let $\tilde{W}_\sigma$ represent a sequence of uncertain payoffs that increase in their standard deviation σ. The greater the σ is, the more uncertainty the payoff $\tilde{W}_\sigma$ represents. This uncertain payoff sequence has a common expected value $E\tilde{W}$. $\tilde{W}_\sigma$ is considered to be a moderate uncertain payoff if $W^{CE}(\tilde{W}_\sigma, z_0, k) \in (s, S)$. When σ is large enough, $W^{CE}(\tilde{W}_\sigma, z_0, k)$ will be outside the range (s, S) and $\tilde{W}_\sigma$ represents a large uncertain payoff. Therefore, the presence of a large uncertain payoff will induce households to abandon their current commitments, while a moderate uncertain payoff will not. When faced with a moderate uncertain payoff, households would rather change freely adjustable consumption goods than pay an adjustment cost to change the consumption of commitment goods.

Chetty and Szeidl (2007) show that commitments amplify the welfare cost of moderate uncertain payoff relative to large uncertain payoff. This result contains two parts. First, for any moderate uncertain payoff denoted by $\tilde{W}$, the higher adjustment cost results in increased risk aversion. To put it differently, households with more commitments, in the sense of higher adjustment costs, require a greater risk premium when faced with a moderate uncertain payoff. Formally, for any $k_1 < k_2$, $\pi(\tilde{W}_\sigma, z_0, k_1) \leq \pi(\tilde{W}_\sigma, z_0, k_2)$. Second, for uncertain payoff sequence $\tilde{W}_\sigma$, equivalent relative risk aversion, $\gamma(\tilde{W}, z_0, k)$ decreases in the size of uncertain payoffs. That is to say, households with commitments demand a larger risk premium per unit for small uncertain payoffs than for large uncertain payoffs. The underlying intuition of this result is moderate uncertain payoffs do not usually induce households to give up commitments, but households with commitments deviate from their optimal consumption plans to adjust to the shocks. The deviation increases in their adjustment costs of commitment.

The result is based on a condition that $W^{CE}(\tilde{W}, z_0, k) < E\tilde{W}$, which indicates that the household's wealth level lies within a concave region of the value function, where the household would like to buy insurance to avoid losses from potential adverse events. The household's purchase of insurance is to avoid having to abandon current commitments after adverse events (negative shocks). This requires that the wealth level after paying the insurance premium is within the (s, S) band. If the insurance premium is so high that the post-purchase wealth is smaller than s, insurance will not be attractive to the household. Therefore, the insurance represents a moderate unexpected payoff with $W^{CE} \in (s, S)$. W^{CE} is the lowest level of post-insurance-purchase wealth required for the household to buy insurance. According to Chetty and Szeidl (2007), households with higher commitment adjustment costs are more likely to buy insurance.

2.3.2 Model Extension: Risk Lovers

This paper extends the Chetty and Szeidl (2007) model to investigate the effect of commitments on risk loving (or risk seeking) behavior, defined as when $W^{CE}(\tilde{W}, z_0, k) > E\tilde{W}$.

When an individual's wealth level is in the neighborhood of $W = s$ or $W = S$, where the value function has a convex curvature, W^{CE} can be greater than $E\tilde{W}$. See Figure 2.2. For this individual, the risk premium is $\pi(\tilde{W}, z_0, k) = 1 + W^{CE}(\tilde{W}, z_0, k)/E\tilde{W}$ and can be interpreted as the minimum payment required for households to give up gambling opportunities. Households have an incentive to both purchase insurance and gamble if their wealth levels are around the transitional region from the concave section to the convex section as the model developed by Friedman and Savage (1948) predicts. To characterize the effect of commitments on the decision to gamble, we need to first define a new uncertain payoff category:

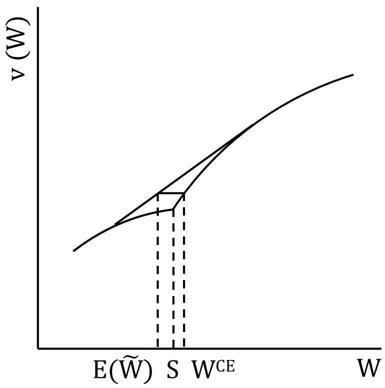

Figure 2.2: Convex Section of Value Function

moderate-to-large (ML) uncertain payoffs. $\tilde{W}$ is a moderate-to-large uncertain payoff if $W^{CE}(\tilde{W}, k_1) \in (S^{k_1}, S^{k_2})$ and $W^{CE}(\tilde{W}, k_2) < S^{k_2}$, when an individual's wealth level is in the neighborhood of $W = S$.[2] Intuitively, a moderate-to-large uncertain payoff is one with certainty equivalent large enough to induce households to abandon their commitments with moderate adjustment costs but not large enough to induce them to abandon commitments with large adjustment costs.

Proposition 1 *Assume the assumptions from Chetty and Szeidl (2007) model hold.*[3] *If* $W^{CE}(\tilde{W}, z_0, k) > E\tilde{W}$, *then*

(1) Adjustment costs mitigate moderate-state risk loving. For any moderate uncertain payoff $\tilde{W}$ *and* $k_1 < k_2$, $\pi(\tilde{W}, z_0, k_1) \geq \pi(\tilde{W}, z_0, k_2)$;

(2) Adjustment costs magnify large-state risk loving. For any large uncertain payoff $\tilde{W}$ *and* $k_1 < k_2$, $\pi(\tilde{W}, z_0, k_1) \leq \pi(\tilde{W}, z_0, k_2)$;

(3) Moderate adjustment costs magnify moderate-to-large-state risk loving while large adjustment costs mitigate ML-state risk loving. For any ML uncertain payoff $\tilde{W}$ *and* $k_1 < k_2$, $\pi(\tilde{W}, z_0, k_1) \leq \pi(\tilde{W}, z_0, k_2)$ *if* $k_2 \leq K$; *and* $\pi(\tilde{W}, z_0, k_1) > \pi(\tilde{W}, z_0, k_2)$ *if* $k_2 > K$. *K is a certain level of adjustment*

[2] For wealth levels in the neighborhood of $W = s$, $\tilde{W}$ is a moderate-to-large uncertain payoff if $W^{CE}(\tilde{W}, k_1) \in (s^{k_2}, s^{k_1})$ and $W^{CE}(\tilde{W}, k_2) > s^{k_2}$.

[3] (A1) Limit properties of utility: $\lim_{g\to\infty} u_1(g, z) = \lim_{z\to\infty} u_2(g, z) = 0$; and $\lim_{g\to 0} u(g, z) = \inf_{g',z'} u(g', z')$ for all z.

(A2) The marginal utility of the freely adjustable good is nondecreasing in housing consumption: $u_{1,2}(g, z) \geq 0$.

(A3) $u(g, z)$ is homogenous of some degree $1 - \gamma$.

(A4) $u(g, z)$ is separable, γ_g is constant, and $\sup_z \gamma_z(z) < \gamma_g$.

cost contingent on k_1 and $K > k_1$.

Proof. First consider the case when the household's wealth level is around S. The household faces an uncertain payoff $\tilde{W}$ with respect to k and z_0. Following Chetty and Szeidl (2007), $k_1 < k_2$ implies $v^{k_1}(W) \geq v^{k_2}(W)$ for all W. See Figure 2.3.

There exists a function $f(W, k_2) \in [0, \bar{k}]$ and $\bar{k}$ defined by $g(k_2 - k_1)$. $f(W, k_2)$ satisfies

$$v^{k_1}(W) = v^{k_2}(W) + f(W, k_2). \tag{2.5}$$

Specifically, $f(W, k_2) = 0$ if $W < S^{k_1}$ and $f(W, k_2) = \bar{k}$ if $W > S^{k_2}$. Because the value function $v(W)$ decreases in adjustment costs k, $f(W)$ increases in k. Taking expectations on both sides of Equation (2.5),

$$E(v^{k_1}(W)) = E(v^{k_2}(W)) + E(f(W, k_2)).$$

Therefore,

$$v^{k_1}(W^{CE}(\tilde{W}, k_1)) = E(v^{k_1}(\tilde{W})) = E(v^{k_2}(\tilde{W})) + E(f(\tilde{W}, k_2)) = v^{k_2}(W^{CE}(\tilde{W}, k_2)) + E(f(\tilde{W}, k_2)). \tag{2.6}$$

Since

$$v^{k_1}(W^{CE}(\tilde{W}, k_1)) = v^{k_2}(W^{CE}(\tilde{W}, k_1)) + f(W^{CE}(\tilde{W}, k_1), k_2);$$
$$v^{k_1}(W^{CE}(\tilde{W}, k_2)) = v^{k_2}(W^{CE}(\tilde{W}, k_2)) + f(W^{CE}(\tilde{W}, k_2), k_2),$$

it follows that

$$v^{k_2}(W^{CE}(\tilde{W}, k_2)) + E(f(\tilde{W}, k_2)) = v^{k_2}(W^{CE}(\tilde{W}, k_1)) + f(W^{CE}(\tilde{W}, k_1), k_2); \tag{2.7}$$
$$v^{k_1}(W^{CE}(\tilde{W}, k_1)) - E(f(\tilde{W}, k_2)) = v^{k_1}(W^{CE}(\tilde{W}, k_2)) - f(W^{CE}(\tilde{W}, k_2), k_2). \tag{2.8}$$

Because $0 \leq E(f(\tilde{W}, k_2)) \leq k$ and $\pi(\tilde{W}, k) = 1 + W^{CE}(\tilde{W}, k)/E\tilde{W}$,

Part (1): if $W^{CE}(\tilde{W}, k_1) < S^{k_1}$, $f(W^{CE}(\tilde{W}, k_1), k_2) = 0$, then $f(W^{CE}(\tilde{W}, k_1), k_2) \leq E(f(\tilde{W}, k_2))$, thus $W^{CE}(\tilde{W}, k_1) \geq W^{CE}(\tilde{W}, k_2)$ and $\pi(\tilde{W}, k_1) \geq \pi(\tilde{W}, k_2)$.

The first part of the proposition demonstrates that households with higher adjustment costs require smaller risk premiums when faced with moderate uncertain payoffs ($W^{CE}(\tilde{W},k) \in (s,S)$). They behave in a less risk-loving manner compared to households with lower adjustment costs.

Part (2): if $W^{CE}(\tilde{W},k_1) > S^{k_1}$ and $W^{CE}(\tilde{W},k_2) > S^{k_2}$, $f(W^{CE}(\tilde{W},k_2),k_2) = \bar{k}$, then

$$f(W^{CE}(\tilde{W},k_2),k_2) \geq E(f(\tilde{W},k_2)),$$

thus $W^{CE}(\tilde{W},k_1) \leq W^{CE}(\tilde{W},k_2)$, and $\pi(\tilde{W},k_1) \leq \pi(\tilde{W},k_2)$.

$\tilde{W}$ is a large uncertain payoff if W^{CE} lies outside the (s,S) band. When faced with large uncertain payoffs, as shown in the second part of the proposition, households with higher adjustment costs demand greater risk premiums. They are more likely to participate in forms of gambling with large uncertain prizes, like lotteries.

Part (3): if $S^{k_1} \leq W^{CE}(\tilde{W},k_1) \leq S^{k_2}$ [4] and $W^{CE}(\tilde{W},k_2) \leq S^{k_2}$, there exists a $k^* \in [k_1,k_2]$ such that $W^{CE}(\tilde{W},k_1) = S^{k^*} \in [S^{k_1},S^{k_2}]$. Therefore,

$$f(W^{CE}(\tilde{W},k_1),k_2) = f(S^{k^*},k_2) = g(k^*-k_1).$$

Since $f(W,k^*) \leq g(k^*-k_1)$, $E(f(\tilde{W},k^*)) \leq f(W,k^*) \leq f(W^{CE}(\tilde{W},k_1),k_2)$. Because $E(f(\tilde{W}))$ increases in k, there exists $\epsilon \geq 0$ such that

$$E(f(\tilde{W},k^*)) \leq E(f(\tilde{W},k^*+\epsilon)) \leq f(W^{CE}(\tilde{W},k_1),k_2)$$

Denote $K = \max(k^*+\epsilon)$ such that $E(f(\tilde{W},K)) = f(W^{CE}(\tilde{W},k_1),k_2)$. When $k_2 \leq K$, $E(f(\tilde{W},k_2)) \leq E(f(\tilde{W},K)) = f(W^{CE}(\tilde{W},k_1),k_2)$, then $W^{CE}(\tilde{W},k_1) \leq W^{CE}(\tilde{W},k_2)$, and $\pi(\tilde{W},k_1) \leq \pi(\tilde{W},k_2)$; when $k_2 > K$, $E(f(\tilde{W},k_2)) > E(f(\tilde{W},K)) = f(W^{CE}(\tilde{W},k_1),k_2)$, thus $W^{CE}(\tilde{W},k_1) > W^{CE}(\tilde{W},k_2)$, and $\pi(\tilde{W},k_1) > \pi(\tilde{W},k_2)$.

The scenario when households are faced with moderate-to-large uncertain payoffs is more complicated. We interpret small prize gambles as casino gambling, slot machines, or video lottery terminal play, whose standard deviation of payoffs are smaller than lotteries, to be a moderate-to-large uncertain payoff. The impact of commitments over moderate-to-large uncertain payoffs exhibits a twofold pattern, depending

[4] It can be shown that if $W^{CE}(\tilde{W},k_1) > S^{k_2}$, then $W^{CE}(\tilde{W},k_2) > S^{k_2}$. This is included in Part (2).

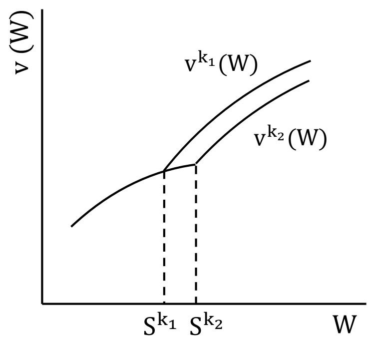

Figure 2.3: Adjustment Cost and Value Function

on the adjustment costs on the commitment goods. Households with moderate commitments and adjustment costs, for example, vehicle-related spending or insurance premium payments, are more likely to participate in small prize gambles; households with higher adjustment costs and commitments, such as a large mortgage parment, are less likely to do so.

The neighborhood of household wealth around s has identical curvature properties as the neighborhood of household wealth around S. The same proof for the case of $W = S$ also applies to the case of $W = s$. ■

What intuition underlies these results? Households with higher adjustment costs have to deviate farther from their optimal consumption plans if they continue with current commitments. In terms of gambling, when uncertain payoffs are moderate, households are not likely to give up their current commitments if they win. Therefore households with higher adjustment costs are less likely to gamble on activities associated with moderate uncertain payoffs and put themselves in a situation of deviating from their optimum consumption path. In contrast, activities associated with large uncertain payoffs would induce households to abandon their current commitments. Households with higher adjustment costs would require higher risk premiums to forgo gambling activities to compensate the cost of maintaining their existing commitments.

A moderate-to-large uncertain payoff is large enough to induce households to abandon existing moderate commitments but not large enough to induce them to abandon large commitments. Therefore, households with moderate adjustment costs may behave in a risk-loving manner in terms of moderate-to-large uncertain payoffs while households with higher adjustment costs avoid participating in gambling activities with moderate-to-large uncertain payoffs. The magnifying or mitigating effect of commitments is size-dependent in this case.

2.4 Empirical Analysis

The model developed in the previous section generates specific predictions about the likelihood an individual simultaneously gambles and purchases insurance when insurance can be interpreted as a commitment good and the form of gambling differs by the size of the uncertain payoff generated. Little empirical evidence addresses actual decisions by individuals to gamble and purchase insurance. This paper next undertakes an empirical analysis of consumer decisions to simultaneously gamble and purchase insurance.

2.4.1 Data Description

Survey of Household Spending

The data used in this paper comes from Statistics Canada's Survey of Household Spending (SHS) from 2004 to 2009. The SHS confidential micro data includes detailed information about household expenditure and other characteristics. In particular, the SHS contains information about household expenditures on government-run lotteries, casino gambling, slot machines and video lottery terminals (VLT), as well as private health insurance and life insurance, together with detailed information on household demographic characteristics.

The SHS includes data collected from residents of all ten Canadian provinces annually since 1997, and the three Canadian territories biannually since 1999. About 20,000 households were interviewed annually before 2007. The sample size was reduced by nearly 30% in 2008 due to Statistics Canada budget cuts. The response rates for the SHS has been about 66%, varying from 63.4% in 2008 to 76.2% in 2001.

This paper analyzes spending decisions made by a relatively homogeneous subsample of the SHS. The analysis sample contains homeowners only. In particular, it consists of single households and couples with children younger than 16, reporting positive household income and paid employment income as their major source of income. Parents with older children and households with other major income sources, such as investments or government transfer payments, are excluded as they may behave quite differently from the households in the sample. Also, households with member(s) older than 65 are excluded.

The sample contains 17,321 observations including annual data from tens of thousands of Canadian households over the period 2004 to 2009. Due to the nature of confidential SHS data, we cannot determine whether the same households participated in the survey in multiple years. Thus, this paper assumes that all 17,321 observations were from different households and treats the data set as repeated cross sections rather than panel data. Summary statistics are shown on Table 2.1. Province and year dummy variables do not appear on this table, but are included in all empirical models. Note that we can not report the maximum and minimum values due to the Statistics Canada confidentiality policy. Instead, this paper reports the values at the 0.1% percentile and 99.9% of the distribution of each variable. These values are very similar to the maximum and minimum values.

Variable Selection

The variable *Government-run Lottery* is annual household spending on government-run lotteries such as Lotto 649 and Lotto Max. On average, about 65% of the households in the sample purchased tickets for these lotteries over the sample period. The variable *Small Prize Gambling* is annual household spending on casinos, slot machines and video lottery terminals. These types of games generally have smaller prizes than lotteries and are thus categorized as moderate-to-large uncertain payoffs in the context of the model developed above. On average, about 18% of the households in the sample spent money on small prize gambling.[5] Over two thirds of households participated in either government-run lottery or small prize gambling.

Besides expenditure on gambling, the SHS contains information on spending on

[5]Ideally, this variable could be disaggregated into several more narrowly-defined spending categories. After all, most casinos have slot machines and/or VLTs, and the difference in payoffs between casino gambling and slot machines or VLTs is larger than that between lotto 649 and lotto max. Unfortunately, the SHS aggregates spending at casinos, on slot machines and on VLTs into a single category.

Table 2.1: Summary Statistics

Variable	Mean	Std. Dev.	0.1st percentile	99.9th percentile
Participation in				
Government-run Lottery	0.6486	0.4774	0	1
Small Prize Gambling	0.1767	0.3814	0	1
Gambling (either type above)	0.6765	0.4678	0	1
Life Insurance	0.5261	0.4993	0	1
Private Health Insurance	0.5511	0.4974	0	1
Insurance (either type of above)	0.7561	0.4295	0	1
Gambling and Life Insurance	0.3780	0.4849	0	1
Gambling and Private Health Insurance	0.3858	0.4868	0	1
Gambling and Insurance	0.5299	0.4991	0	1
Age	43.5593	10.6448	21	65
Income (0000)	7.8214	7.3511	0.5100	58.6618
Metropolitan	0.4786	0.4996	0	1
Number of Vehicles	1.6228	0.7270	0	5
Housing Tenure				
Homeowner without a mortgage	0.2896	0.4536	0	1
Homeowner with mortgage	0.7104	0.4536	0	1
Household Type				
Single female	0.0906	0.2871	0	1
Single male	0.0998	0.2997	0	1
Couple only	0.3851	0.4866	0	1
Couple with children	0.4245	0.4943	0	1
Household Employment Status				
At least one person working part time	0.5500	0.4975	0	1
All person(s) working full time	0.4500	0.4975	0	1
Household Highest Level of Education				
No degrees, certificates or diplomas	0.0551	0.2281	0	1
High school or equivalent	0.1752	0.3801	0	1
Certificate or diploma below Bachelor's	0.4396	0.4964	0	1
Bachelor's degree	0.2046	0.4034	0	1
Certificate or diploma above Bachelor's	0.1255	0.3313	0	1

several forms of insurance, including tenants' or homeowners' insurance, life insurance, employment insurance, vehicle insurance and health care insurance. Most of these forms of insurance are mandatory. For example, many tenants must agree to purchase insurance when they sign a lease on an apartment or home; homeowners with mortgages are required by lenders to have homeowners' insurance. Employment insurance is deducted directly from pay cheques, and vehicle insurance is required by law. There are only two forms of voluntary insurance in the SHS: life insurance and private health care insurance. The variable *Private Health Care Insurance* is expenditure on private health insurance plans, which include supplementary coverage to publicly funded hospital and medical plans, extended health benefit packages, drug plans, out-of-country benefits and visitors' benefits, dental plans, and accident and disability insurance. About 55% of the households in the sample purchased private health care insurance in the sample period. We should however point out that those who did not purchase such insurance in this sample might have already been insured through their employers' insurance plans. Those who did purchase such insurance in our data, on the other hand, might have purchased additional extra insurance policies above those provided by employers. *Life Insurance* is the premium a household paid for life, term and endowment insurance policies. About 53% of the households in the sample purchased life insurance. Life insurance basically covers "end of life expenses." The insured individuals' objective is to support their loved ones after their death and not leave any financial responsibilities. Compared to spending on *Private Health Care Insurance*, *Life Insurance* is more likely to reflect individuals' preference and choices because they are less likely to have been covered through their employers, who tend to provide only basic life insurance, if at all.

The summary statistics in Table 2.1 also show that over half of the sample households have both purchased voluntary insurance (life insurance and/or private health care insurance) and participated in gambling activities. About 38% of the households spent money on life insurance and at least one of the two types of gambling. That means almost 72% of the households who purchased life insurance also bought lottery tickets and/or participated in some form of small prize gambling. The simultaneous purchasing of insurance and gambling is a common behavior in this sample.

Income is the household's real total annual income before taxes in the reference year. The average income in the sample is close to $78,000; the top earning households in the sample earned over half million. *Age* is the age of reference person in the reference year. For single households, the age of reference person is the age of household. For couple households, the age of reference person is usually close to the age of his/her spouse. Therefore the reference person's age is used as the age of the couple household. The average age is 44. *Number of Vehicles* is the number of vehicles a household owned or leased in the reference year. On average, households in the sample have 1.6 vehicles.

The rest of the variables on Table 2.1 are all indicator variables. *Metropolitan* identifies whether the household lives in a metropolitan area. Statistics Canada's definition of metropolitan area is "a very large urban area (known as the urban core) together with adjacent urban and rural areas that have a high degree of social and economic integration with the urban core." Nearly half of the households in the sample reside in metropolitan areas. Other variables are categorized in 4 groups: *Housing Tenure Group*, *Household Type*, *Household Employment Status* and *Household Highest Level of Education*.

There are 2 housing tenure types in the sample. Around 71% of the households are homeowners who have mortgages. Homeowners without a mortgage account for the rest less than 30% of the sample. Households that reported to both own and rent dwellings in the same year are excluded from the sample. We expect people with more commitments are more likely to gamble and purchase insurance. The participation rates showed in Table 2.2 are consistent with this conjecture. About 68% and 53% of homeowners with mortgages engaged in gambling activities and life insurance purchase, respectively, 2% and 3% more than homeowners who already paid off their mortgages.

Table 2.2: Participation Rate for Homeowners

	Homeowner w/o a mortgage	Homeowner w/ mortgage
Gambling	66.18%	68.25%
Government-run Lottery	63.86%	65.27%
Small Prize Gambling	16.64%	18.08%
Life Insurance	50.37%	53.52%

There are more couple households than single households. Of the 19% one-person households, there are more single males, by about 10%, than single females. *Household Employment Status* consists of 2 variables identifying 2 different labour market status: (1) at least one person working part time; and (2) all person(s) working full time. In 55% of the households in the sample, at least one person was working during the reference year, but none of the household member(s) had a full time job. Finally, there is a set of dummy variables that identify 5 levels of educational attainment, from "no degree" to "above Bachelor's." The largest group, about 44%, is households that have certificates or diplomas below Bachelor's degrees. About 33% of the households have a Bachelor's degree or above, while households with no degree, certificate or diploma account for less than 6% of the total. Households that are missing the education information are excluded.

2.4.2 Empirical Approach

This paper develops evidence about the causal connection between participation in different types of gambling, the purchase of insurance, and household characteristics including consumption commitments. The empirical analysis poses an econometric challenge because these purchase decisions are unlikely to be independent. Factors that affect spending on one form of gambling may affect spending on another type of gambling. Factors that affect how much insurance to purchase can also affect spending on gambling. Some of those factors can be observed and thus controlled for in the regression analysis. Those that are not directly observed, such as past experience and individual risk assessment, will fall into the error terms of the regression models. These error terms are very likely correlated with each other.

When interdependent error terms exist across multiple regression equations, a seemingly unrelated regressions (SUR) model is often used. The SUR model assumes that error terms are uncorrelated across observations but correlated across equations. In addition, this paper analyzes the impact of consumption commitments on households' decisions on gambling and insurance purchase, which are measured as binary outcome variables. One special case of the SUR model, the multivariate probit model, has been widely used in applications where binary dependent variables are simultaneously determined. Hyslop (1999) applies this technique to investigate the labor force participation decision of women using longitudinal data. Other examples in the literature include Greene (2004) and Cappellari and Jenkins (2004).

This paper follows Greene (2008) and uses a general form of an M-variate probit model, analogous to the SUR model,

$$\begin{aligned}
&y_m^* = x_m'\beta_m + \varepsilon_m, \\
&y_m = 1 \quad \text{if} \quad y_m^* > 0, \quad 0 \quad \text{otherwise}, \quad m = 1, ..., M, \\
&E[\varepsilon_m \mid x_1, ..., x_M] = 0, \\
&Var[\varepsilon_m \mid x_1, ..., x_M] = 1, \\
&Cov[\varepsilon_j, \varepsilon_m \mid x_1, ..., x_M] = \rho_{jm}, \\
&(\varepsilon_1, ..., \varepsilon_m) \sim N_M[0, R].
\end{aligned} \tag{2.9}$$

In this setting, the log-likelihood function for the normal distribution is

$$\ln L = \sum_{y_m=0} \ln[1 - \Phi(x_m'\beta)] + \sum_{y_m=1} \ln \Phi(x_m'\beta). \tag{2.10}$$

The joint probabilities are

$$\begin{aligned}
&L_i = \Phi_M(q_{i1}x_{i1}'\beta_1, ..., q_{iM}x_{iM}'\beta_M; R^*), \\
&q_{im} = 2y_{im} - 1, \\
&R_{jm}^* = q_{ij}q_{im}\rho_{jm}.
\end{aligned}$$

If $\rho = 0$, the multivariate probit model becomes an M-independent univariate probit model. The log-likelihood for the multivariate probit model is then equal to the sum of the log-likelihoods for each univariate probit model. A likelihood ratio test can be performed by comparing the overall likelihood and the sum of the M separate log-likelihoods. ρ is estimated indirectly through an estimator of $\arg\tanh\rho$

$$\arg\tanh\rho = \frac{1}{2}\ln(\frac{1+\rho}{1-\rho}).$$

This parameter can also be estimated directly using maximum likelihood estimation.

In this study, we first estimate positive spending on *Gambling* (either government-run lottery or small prize gambling) and *Life Insurance* simultaneously, so $M = 2$ and we have a bivariate probit model. In the second step, we divide *Gambling* into *Government-run Lottery* and *Small Prize Gambling*, and estimate positive spending on these two gambling activities and *Life Insurance* simultaneously, then $M = 3$ and we have a trivariate probit model.

The log-likelihood function for the more complex trivariate probit model is

$$\ln L = \sum_{i=1}^{N} \log \Phi_3(q_{i1}x'_{i1}\beta_1, q_{i2}x'_{i2}\beta_2, q_{i3}x'_{i3}\beta_3; R^*) \tag{2.11}$$

where N is the sample size and Φ_3 is the cumulative distribution function (cdf) of the trivariate normal distribution, where

$$R^*_{jj} = 1 \quad for \;\; j = 1, ..., 3$$
$$R^*_{jm} = R^*_{mj}$$

The multivariate probit model is a nonlinear regression model. Its parameters cannot be interpreted as marginal effects. The marginal effect for a trivariate probit model can be evaluated by differentiating Φ_3 with respect to x_j. If the independent variable is a dummy variable, this derivative approach is not appropriate. For dummy variables, the marginal effect is the simple difference found by comparing the trivariate probability at the two values of the dummy variable d,

$$\text{Marginal effect} = \Phi_3(X, R^*|d = 1) - \Phi_3(X, R^*|d = 0).$$

As Greene (2008) noted, the practical difficulty when implementing the multivariate probit regression model is "the evaluation of the M-variate normal integrals and their derivatives." Many studies use the GHK simulator to perform simulation-based integration,[6] such as Greene (2004), Cappellari and Jenkins (2004) and Hyslop (1999). The simulation-based estimation of maximum-likelihood functions is computationally intensive, especially in large samples like this. To facilitate estimation, some studies such as Greene (2004) used pseudorandom draws from a uniform distribution, specifically Halton sequences, instead of the more conventional random draws from a standard continuous uniform distribution.[7]

Several papers have investigated the relationship between expenditure on gambling, especially lottery, and individual's economic and demographic characteristics.

[6]See Hajivassiliou et al. (1996) and Geweke et al. (1997) for details about the GHK simulator.

[7]See Greene (2004) for details about Halton sequences.

Scott and Garen (1994), Abdel-Ghany and Sharpe (2001), Worthington et al. (2007) and Humphreys and Perez (2012) study consumers' expenditure on lottery controlling for some economic and demographic characteristics including income, occupation, age, sex, race, education and household structure. Besides lottery tickets, Worthington et al. (2007) also investigate household spending on three other types of gambling, including slot machines, using Tobit models. All the examples above use micro-level survey data. Among them, Abdel-Ghany and Sharpe (2001) and Worthington et al. (2007) use household spending surveys similar to the SHS.

The vector of covariates x_m includes age, age squared, income, income squared, number of vehicles owned or leased, and the dummy variables discussed above. Year and province indicators are also included in the regression models. Since insurance premium payments are consumption commitments, we include *Private Health Insurance* and *Life Insurance* in x_m to reflect this spending. Since life insurance is on the right-hand side of the gambling participation equations, the equation system is a recursive multivariate probit model. Wilde (2000) shows a recursive multivariate probit model is identified if each x_m contains one varying exogenous variable, which is easily satisfied in this model. Greene (2008) also mentions the endogenous nature of one dependent variable on the right-hand side "can be ignored in formulating the log-likelihood" when estimating a recursive multivariate probit model (page 823).

Finally, we note that the survey design of the SHS involves sampling households from Statistics Canada defined geographical areas called "*Clusters*." Since residents of a cluster may share common unobservable social and economic characteristics, there is a potential for intra-cluster correlation in observed as well as unobservable factors influencing insurance and gambling decisions. We use cluster-robust standard errors at the level of *Cluster* to correct for the potential biases from heteroskedasticity and intra-cluster correlation in the equation error terms.

2.4.3 Results and Discussion

The marginal effects, standard errors and p-values obtained from estimating the bivariate and trivariate probit models (with and without cluster-corrected standard errors) are shown on Tables 2.4-2.9. Non-corrected standard errors and p-values are reported because clustered standard errors would be biased upward from the true value if no within-cluster correlation exists. However, results that are not cluster-corrected are broadly similar to those cluster-corrected. We do not report marginal effects for year and province indicator variables since their effects are not of primary interest here. Most of ρ's are significantly different from 0 in those multivariate probit models.

Tables 2.4 and 2.5 show the results for bivariate probit models. Each table contains estimates from two models: one includes variables indicating insurance purchase (right panel) and the other does not (left panel). The results are quite similar in terms of estimated parameter signs and significance across two models. However, including insurance variables reduces other estimates by roughly 55-67 percent.

The parameter estimates of interest are those on the upper half of the table, which shows results for variables reflecting consumption commitments. Residential housing and household size are often viewed as spending commitments (e.g. Calvet and Sodini (2014), Chetty and Szeidl (2007)). We proxy for the presence of a large consumption commitment with a home mortgage. A homeowner with a mortgage will pay a high adjustment cost if she decides to move. For example, she will likely have to pay a real estate agent to sell the property, and is responsible for monthly mortgage, utility fees and other expenses, as long as the property is on the market. If the property is on the market for several months (which is common), it will result in a high adjustment cost, mostly from mortgage payments. Household types are also used to indicate the level of consumption commitments. Larger households usually have more consumption commitments than smaller households. Housing tenure group *homeowner without a mortgage* and household type *single female* are omitted because of collinearity during estimation.

The following discussion focuses on the results on the right panel of Table 2.5, which includes insurance variables and shows marginal effects and corrects for potential intra-cluster correlations and heteroskedasticity. As predicted in the theoretical model, since consumption commitments change the shape of the value function and induce local non-concavities, households with commitments are more likely to both gamble and purchase insurance. Homeowners with mortgages are about 6% more likely to gamble and 3% more likely to purchase life insurance than homeowners without a mortgage, holding income, employment status, age, eduction and other factors constant. The effect of an additional household member on the likelihood of purchasing life insurance is substantial. Compared to single female households, couples are about 11% more likely to purchase life insurance. Couple-with-children households are nearly 20% more likely to do so. As for gambling activities, couples are almost 7% more likely to participate in than single female. However, couple-with-children households are not more likely to gamble than single female households, probably due to parents' concerns about potential negative effects of gambling on kids.

Other commitment variables tell a similar story: households who have more commitments, thus bearing higher adjustment costs, are more likely to engage in both gambling and insurance purchase. Owning or leasing more vehicles, which is also a form of commitment as it involves necessary spending on car insurance, monthly financing or lease payments, gas, maintenance and repair, also increases gambling participation and insurance purchase. Specifically, one more vehicle in the household increases the probabilities of gambling and life insurance purchasing by about 4% and 2%, respectively. Households who purchase private health insurance are about 4% more likely to gamble and 8% more likely to buy life insurance. Not only is it possible that households purchase insurance and gamble simultaneously, but insurance purchasing as a commitment increases the probability of households' participating in gambling activities. The insignificance of parameter estimate of *life insurance* might be due to some multicollinearity between *private health care insurance* and *life insurance*.

Variables on the lower half of Table 2.5 are control variables reflecting households' demographic characteristics. Those who work full time are more likely to purchase insurance and to gamble. This may reflect decisions consistent with the model developed above. Consider time as an endowment and work as a commitment. In this context, adjustment costs, which include loss of earnings, are higher for full time workers than for part time workers and the model would predict that full time workers are more likely to both buy life insurance and gamble. Living in a metropolitan area has negative but insignificant effect on households' purchase of both lottery tickets and life insurance. The inclusion of age squared and income squared allows for a non-linear relationship between participation in gambling and purchase of life insurance, and both age and income. In all three equations, the estimated parameters on age,

age squared, income and income squared are significantly different from zero. The relationship between age and gambling participation takes the standard hump-shaped form, which is consistent with the findings in Farrell and Walker (1999) and Clotfelter and Cook (1991). Higher educational attainment increases the purchase likelihood for life insurance, with those with Bachelor's degrees most likely to buy life insurance. However, for gambling activities, only households with education attainment above a Bachelor's degree are significantly less likely to gamble compared to households with no degree, certificate or diploma. Those who have lower levels of education behave similarly to "no degree" households.

The theoretical model in last section suggests that consumption commitments have different effects on household's risk preference when facing different size of uncertain payoffs. The trivariate probit model is employed to investigate effect of commitments on likelihood for participation in activities with large uncertain payoffs (e.g. lottery ticket purchase), activities with moderate-to-large uncertain payoffs (e.g. small prize gambling), and activities with moderate uncertain payoffs (e.g. insurance purchase). Table 2.6 shows parameter estimates for the trivariate probit model with the same variables in the first bivariate probit model. The results are largely consistent with the prediction of the model developed above. Households with more consumption commitments associated with higher adjustment costs are more likely to purchase government-run lottery ticket and life insurance. However, a home mortgage seems not to have an impact on likelihood for participation in small prize gambling, which includes casino gambling, slot machines and VLTs. Because effect of commitments on those gambling activities with moderate-to-large uncertain payoffs is size-dependent, we split the mortgage payment into four groups and use four mortgage size indicators instead of a single indicator reflecting having mortgage or not. The quartiles are 0, 6720, and 11340. Since 29% of households have no mortgage payment, the second group, in which households pay mortgage less than 6720 dollars in the reference year, contains about 21% of the households in the sample. The last two groups have equally 25% of households.

The results from estimating trivariate probit models including four mortgage size indicators (*mortgage quarter 1-4*) are shown on Tables 2.7-2.9. *Mortgage quarter 1* is omitted due to collinearity during estimation. Models with and without cluster-corrected standard errors and insurance variables generate quite similar results in terms of estimate signs, significance, and magnitude. Estimated parameters for non-mortgage related variables are also broadly similar across models with four mortgage size variables and models with a single mortgage indicator. Our discussion of mortgage size effect focuses on estimates reported on Table 2.9, which includes insurance variables and corrects for potential intra-cluster correlations. As predicted in the theoretical model, higher mortgage payment increases the purchasing likelihood for lottery ticket and life insurance. Further, as the mortgage payment increases, the probability of participation in small prize gambling increases in the beginning and falls afterwards. The findings about small prize gambling support the theoretical prediction that moderate adjustment costs magnify ML-stake risk loving while larger adjustment costs mitigate ML-stake risk loving.

We now move on to effects of other explanatory variables. The parameter estimates shown on Table 2.9 suggest that couple households are more likely to engage in gambling with large uncertain payoffs, namely government-run lottery than single households. However, couples are not more likely to participate in gambling with moderate-to-large uncertain payoffs, like casino gambling than singles. In particular, couple-with-children households are about 9% less likely to gamble in casinos or with slot machines/VLTs. Unlike in the bivariate probit model, purchase of life insurance increases the probability of purchasing government-run lottery ticket by more than 7%, as well as the likelihood of engaging in small prize gambling by almost 2%. Full time working households are more likely to buy lottery tickets, but not more likely to participate in small prize gambling activities. The estimated relationship between employment and purchase of government-run lottery tickets is consistent with the predictions in the model developed by Nyman et al. (2008) about the relationship between employment status and lottery participation.

2.4.4 Robustness Checks

Discussions above are based on results from estimations using an aggregate sample including four different household types: single female, single male, couple only, and couple with children. However, risk preferences and purchase behaviors may be different across household types. We performed a series of robustness checks on the results reported on Tables 2.5 and 2.9 using subsamples containing only one type of households. The marginal effects, standard errors and p-values obtained from those estimations are shown on Tables 2.10-2.17. The discussion here focuses on the estimated parameter on the most interesting variable, home mortgage. For couple-with-children households, the results are largely consistent with those obtained from estimating the whole sample. For couple-only households, mortgages increase the likelihood of gambling participation but have no significant impact on life insurance purchase. When only investigating spending on gambling activities as a whole and life insurance, the parameter estimates suggest that single female households with mortgages are more likely to gamble, while single males who pay mortgages are more likely to purchase life insurance. If we divide gambling activities into lottery purchase and small prize gambling, moderate mortgage payment increases single female homeowners' likelihood of participation in small prize gambling and larger mortgage payment increases their probability of purchasing lottery tickets. However, home mortgages only increase single male homeowners' purchasing likelihood for lottery ticket, but not for small prize gambling.

Since the survey is conducted annually, households may not remember if they had the spending on gambling or insurance in the reference year at the time of interview, especially when the spending was very small. Some households might record some small numbers in the interview just in case. Another case worth mentioning is that some individuals buy lottery tickets rarely and impulsively. This is a non-typical behavior which is difficult to explain using economic models based on rationality assumption. A common treatment for these cases is looking for the natural gap between pools of spending and recoding the pool near zero as zero, i.e. assuming there was no very small spending. The results of these robustness checks are shown on Tables 2.18 and 2.19. The results, specifically the sign and significance of parameter estimates, are not sensitive to these changes to the data.

2.5 Conclusions

Since Friedman and Savage (1948), local non-concavities in the utility function have been used to explain seemingly contradictory but common phenomenon of individuals simultaneously buying insurance and gambling. Chetty and Szeidl (2007), along with others, offer a new explanation for this local non-concavity. They argue that consumption commitments magnify risk aversion over moderate risks and induce local non-concavities in the value function over wealth. If wealth level is near the transitional region from the concave part to the convex part of the wealth function, the household would both purchase insurance and gamble.

This study extends the model developed in Chetty and Szeidl (2007) by investigating the impacts of consumption commitments on risk loving over activities with large and moderate-to-large uncertain payoffs. We show that commitments increase risk loving over large uncertain payoffs but for activities with moderate-to-large uncertain payoffs, the impacts of consumption commitments are size-dependent: while moderate consumption commitments amplify risk loving, large consumption commitments mitigate risk loving. The magnitude of uncertain payoffs is defined according to their standard deviation. We assume lotteries can be interpreted as large uncertain payoffs while casino gambling, slot machines and video lottery terminals represent moderate-to-large uncertain payoffs.

This paper enhances understanding of the relationship between gambling and insurance. Since insurance can be considered a moderate consumption commitment, people who purchased insurance are more likely to gamble. Not only is it common for people to purchase insurance and gamble at the same time in this sample, but insurance purchase as a consumption commitment increases the likelihood of participation in gambling activities.

This paper also makes an empirical contribution to the literature. Using confidential micro data from the Canadian SHS, the paper simultaneously estimates probit models of household participation in government-run lottery and small prize gambling and their purchases of life insurance. Based on estimates from multivariate probit models controlling for households' economic and demographic characteristics, we find empirical evidence consistent with the predictions of the model. Households with greater consumption commitments, in the forms of a home mortgage, car ownership, etc. are found to be more likely to gamble on lotteries and buy life insurance. For small prize gambling, however, the probability of participation is higher for households who have moderate commitments, but less likely for households with large commitments, as the theory predicts.

This study extends the model developed by Chetty and Szeidl (2007), which is consistent with the standard rational choice approach to consumer choice and includes diminishing marginal utility of consumption. It also contributes to the literature on expected utility theory with local non-concavities. Some clear extensions to this line of research exist. Chetty and Szeidl (2007) interpret unemployment as a shock. However, it would be interesting to incorporate individual work life choice as modelled by Dowell and McLaren (1986) into the model in future work. Full time work can be interpreted as a commitment if time enters the utility function. In this case, the adjustment cost is generated by the loss of earnings.

Table 2.3: Demographic Variable Descriptions

Variable	Abbreviation on Table 2.4-2.19
Household Highest Level of Education	
No degrees, certificates or diplomas	(*omitted*)
High school or equivalent	$HLE2$
Certificate or diploma below Bachelor's	$HLE3$
Bachelor's degree	$HLE4$
Certificate or diploma above Bachelor's	$HLE5$

Table 2.4: Marginal Effects for Bivariate Probit Model with No Cluster Correction

	Gambling			Life Insurance			Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Homeowner w/ mtg.	0.16850	0.02531	<0.001	0.09011	0.02461	<0.001	0.05620	0.00964	<0.001	0.03006	0.00908	0.001
Single male	0.02068	0.04585	0.652	-0.03261	0.04620	0.480	0.00905	0.01579	0.566	-0.00849	0.01706	0.619
Couple only	0.20169	0.04111	<0.001	0.30976	0.04027	<0.001	0.06801	0.02311	0.003	0.11200	0.01478	<0.001
Couple w/ children	0.06407	0.04420	0.147	0.51217	0.04338	<0.001	0.02005	0.03514	0.568	0.18577	0.01582	<0.001
Number of vehicle	0.12529	0.01568	<0.001	0.07334	0.01502	<0.001	0.04193	0.00643	<0.001	0.02482	0.00553	<0.001
Pri. health insurance							0.03915	0.01524	0.010	0.08119	0.00751	<0.001
Life insurance							0.00196	0.16879	0.991			
Working full time	0.08254	0.02290	<0.001	0.07073	0.02205	0.001	0.02722	0.00860	0.002	0.02352	0.00812	0.004
Metropolitan	-0.04696	0.02211	0.034	-0.04400	0.02144	0.040	-0.01601	0.00799	0.045	-0.01623	0.00791	0.040
Age	0.05830	0.00815	<0.001	0.04863	0.00801	<0.001	0.01983	0.00385	<0.001	0.01767	0.00295	<0.001
Age^2	-0.00055	0.00009	<0.001	-0.00040	0.00009	<0.001	-0.00019	0.00004	<0.001	-0.00015	0.00003	<0.001
Income	0.01448	0.00351	<0.001	0.02311	0.00241	<0.001	0.00477	0.00184	0.010	0.00836	0.00088	<0.001
$Income^2$	-0.00011	0.00004	0.012	-0.00004	0.00001	<0.001	-0.00004	0.00002	0.019	-0.00001	0.00000	<0.001
HLE2	0.07952	0.04976	0.110	0.08961	0.04889	0.067	0.02402	0.01756	0.171	0.02677	0.01806	0.138
HLE3	0.12029	0.04697	0.010	0.22659	0.04616	<0.001	0.03607	0.02006	0.072	0.07358	0.01704	<0.001
HLE4	-0.05781	0.05130	0.260	0.32183	0.05036	<0.001	-0.02557	0.02567	0.319	0.10754	0.01857	<0.001
HLE5	-0.36658	0.05500	<0.001	0.25739	0.05413	<0.001	-0.13228	0.02479	<0.001	0.08263	0.01999	<0.001
Observations	17321											

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation $\textit{Marginal effect of age} + 2\overline{Age}\ \textit{Marginal effect of } Age^2$. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.5: Marginal Effects for Bivariate Probit Model with Cluster Corrected Standard Errors

	Gambling			Life Insurance			Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Homeowner w/ mtg.	0.16850	0.02211	<0.001	0.09011	0.02557	<0.001	0.05620	0.00849	<0.001	0.03006	0.00939	0.001
Single male	0.02068	0.04206	0.623	-0.03261	0.03967	0.411	0.00905	0.01426	0.526	-0.00849	0.01494	0.570
Couple only	0.20169	0.04049	<0.001	0.30976	0.04135	<0.001	0.06801	0.02253	0.003	0.11200	0.01510	<0.001
Couple w/ children	0.06407	0.04680	0.171	0.51217	0.04164	<0.001	0.02005	0.03313	0.545	0.18577	0.01552	<0.001
Number of vehicle	0.12529	0.01723	<0.001	0.07334	0.01264	<0.001	0.04193	0.00726	<0.001	0.02482	0.00459	<0.001
Pri. health insurance							0.03915	0.01285	0.002	0.08119	0.00789	<0.001
Life insurance							0.00196	0.15917	0.990			
Working full time	0.08254	0.02521	0.001	0.07073	0.02257	0.002	0.02722	0.00897	0.002	0.02352	0.00823	0.004
Metropolitan	-0.04696	0.02771	0.090	-0.04400	0.02310	0.057	-0.01601	0.01000	0.109	-0.01623	0.00844	0.055
Age	0.05830	0.00817	<0.001	0.04863	0.00773	<0.001	0.01983	0.00402	<0.001	0.01767	0.00288	<0.001
Age^2	-0.00055	0.00010	<0.001	-0.00040	0.00009	<0.001	-0.00019	0.00004	<0.001	-0.00015	0.00003	<0.001
Income	0.01448	0.00399	<0.001	0.02311	0.00463	<0.001	0.00477	0.00196	0.015	0.00836	0.00168	<0.001
$Income^2$	-0.00011	0.00005	0.040	-0.00004	0.00001	<0.001	-0.00004	0.00002	0.046	-0.00001	0.00000	<0.001
HLE2	0.07952	0.05217	0.127	0.08961	0.04507	0.047	0.02402	0.01837	0.191	0.02677	0.01668	0.109
HLE3	0.12029	0.05600	0.032	0.22659	0.04898	<0.001	0.03607	0.02291	0.115	0.07358	0.01792	<0.001
HLE4	-0.05781	0.05574	0.300	0.32183	0.05121	<0.001	-0.02557	0.02689	0.342	0.10754	0.01855	<0.001
HLE5	-0.36658	0.05528	<0.001	0.25739	0.05308	<0.001	-0.13228	0.02527	<0.001	0.08263	0.01947	<0.001
Observations	17321											

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation $\textit{Marginal effect of age} + 2\overline{Age}\ \textit{Marginal effect of } Age^2$. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.6: Marginal Effects for Trivariate Probit Model with No Cluster Correction

	Government-run Lottery			Small Prize Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Homeowner w/ mtg.	0.05766	0.00886	$<$0.001	0.01283	0.00715	0.073	0.03320	0.00910	$<$0.001
Single male	0.01115	0.01614	0.490	-0.01860	0.01312	0.156	-0.01186	0.01710	0.488
Couple only	0.08366	0.01441	$<$0.001	-0.00283	0.01161	0.807	0.11476	0.01483	$<$0.001
Couple w/ children	0.05250	0.01552	0.001	-0.08545	0.01258	$<$0.001	0.18974	0.01585	$<$0.001
Number of vehicle	0.04121	0.00546	$<$0.001	0.02107	0.00423	$<$0.001	0.02714	0.00555	$<$0.001
Working full time	0.02892	0.00801	$<$0.001	0.01189	0.00632	0.060	0.02621	0.00815	0.001
Metropolitan	-0.01818	0.00775	0.019	0.00003	0.00615	0.997	-0.01623	0.00793	0.041
Age	0.02396	0.00285	$<$0.001	-0.00993	0.00224	$<$0.001	0.01797	0.00296	$<$0.001
Age^2	-0.00022	0.00003	$<$0.001	0.00009	0.00003	0.001	-0.00015	0.00003	$<$0.001
Income	0.00367	0.00123	0.003	0.00803	0.00094	$<$0.001	0.00857	0.00089	$<$0.001
$Income^2$	-0.00004	0.00002	0.021	-0.00005	0.00001	$<$0.001	-0.00001	0.00000	$<$0.001
HLE2	0.03100	0.01748	0.076	0.04516	0.01485	0.002	0.03286	0.01808	0.069
HLE3	0.04571	0.01649	0.006	0.03342	0.01419	0.018	0.08354	0.01704	$<$0.001
HLE4	-0.01326	0.01803	0.462	-0.00458	0.01542	0.766	0.11881	0.01857	$<$0.001
HLE5	-0.12175	0.01929	$<$0.001	-0.05073	0.01677	0.002	0.09482	0.01999	$<$0.001
Observations	17321								

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation $\textit{Marginal effect of age} + 2\overline{Age}\ \textit{Marginal effect of Age}^2$. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.7: Marginal Effects for Trivariate Probit Model with No Cluster Correction: Effects of Mortgage Size I

	Government-run Lottery			Small Prize Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Mortgage quarter 2	0.05125	0.01095	<0.001	0.00203	0.00893	0.821	0.01183	0.01122	0.291
Mortgage quarter 3	0.05845	0.01062	<0.001	0.02541	0.00845	0.003	0.04078	0.01089	<0.001
Mortgage quarter 4	0.06576	0.01133	<0.001	0.01135	0.00901	0.208	0.04275	0.01158	<0.001
Single male	0.01103	0.01614	0.494	-0.01875	0.01311	0.153	-0.01287	0.01710	0.452
Couple only	0.08274	0.01443	<0.001	-0.00386	0.01163	0.740	0.11174	0.01487	<0.001
Couple w/ children	0.05102	0.01557	0.001	-0.08690	0.01262	<0.001	0.18561	0.01594	<0.001
Number of vehicle	0.04126	0.00546	<0.001	0.02101	0.00423	<0.001	0.02691	0.00554	<0.001
Working full time	0.02855	0.00802	<0.001	0.01119	0.00632	0.077	0.02493	0.00817	0.002
Metropolitan	-0.01964	0.00783	0.012	-0.00086	0.00620	0.890	-0.01916	0.00802	0.017
Age	0.02390	0.00285	<0.001	-0.00997	0.00224	<0.001	0.01787	0.00296	<0.001
Age^2	-0.00022	0.00003	<0.001	0.00009	0.00003	0.001	-0.00015	0.00003	<0.001
Income	0.00337	0.00125	0.007	0.00800	0.00096	<0.001	0.00830	0.00090	<0.001
Income^2	-0.00003	0.00002	0.035	-0.00005	0.00001	<0.001	-0.00001	0.00000	<0.001
HLE2	0.03140	0.01748	0.072	0.04468	0.01486	0.003	0.03292	0.01808	0.069
HLE3	0.04579	0.01650	0.006	0.03232	0.01420	0.023	0.08284	0.01705	<0.001
HLE4	-0.01392	0.01805	0.441	-0.00593	0.01543	0.701	0.11626	0.01859	<0.001
HLE5	-0.12232	0.01930	<0.001	-0.05205	0.01678	0.002	0.09216	0.02001	<0.001
Observations	17321								

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation $\textit{Marginal effect of age} + 2\overline{Age}\ \textit{Marginal effect of Age}^2$. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.8: Marginal Effects for Trivariate Probit Model with No Cluster Correction: Effects of Mortgage Size II

	Government-run Lottery			Small Prize Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Mortgage quarter 2	0.04859	0.01091	<0.001	0.00154	0.00893	0.863	0.00920	0.01118	0.411
Mortgage quarter 3	0.05366	0.01058	<0.001	0.02455	0.00845	0.004	0.03695	0.01086	0.001
Mortgage quarter 4	0.06242	0.01129	<0.001	0.01066	0.00901	0.237	0.03999	0.01154	0.001
Single male	0.01376	0.01608	0.392	-0.01800	0.01311	0.170	-0.00926	0.01705	0.587
Couple only	0.07487	0.01438	<0.001	-0.00585	0.01163	0.615	0.10915	0.01482	<0.001
Couple w/ children	0.03727	0.01551	0.016	-0.09037	0.01263	<0.001	0.18193	0.01589	<0.001
Number of vehicle	0.03866	0.00545	<0.001	0.02036	0.00423	<0.001	0.02457	0.00553	<0.001
Pri. health insurance	0.03397	0.00740	<0.001	0.00906	0.00593	0.127	0.08117	0.00751	<0.001
Life insurance	0.07229	0.00034	<0.001	0.01716	0.00018	<0.001			
Working full time	0.02653	0.00800	0.001	0.01072	0.00632	0.090	0.02227	0.00814	0.006
Metropolitan	-0.01797	0.00780	0.021	-0.00057	0.00620	0.927	-0.01909	0.00799	0.017
Age	0.02258	0.00284	<0.001	-0.01028	0.00224	<0.001	0.01755	0.00295	<0.001
Age^2	-0.00021	0.00003	<0.001	0.00009	0.00003	<0.001	-0.00015	0.00003	<0.001
Income	0.00205	0.00126	0.103	0.00773	0.00096	<0.001	0.00809	0.00089	<0.001
$Income^2$	-0.00002	0.00002	0.173	-0.00005	0.00001	<0.001	-0.00001	0.00000	<0.001
HLE2	0.02629	0.01744	0.132	0.04371	0.01487	0.003	0.02670	0.01804	0.139
HLE3	0.03552	0.01648	0.031	0.03020	0.01422	0.034	0.07276	0.01703	<0.001
HLE4	-0.02641	0.01801	0.143	-0.00862	0.01545	0.577	0.10487	0.01857	<0.001
HLE5	-0.13332	0.01926	<0.001	-0.05434	0.01680	0.001	0.07976	0.01999	<0.001
Observations	17321								

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation *Marginal effect of age* + $2\overline{Age}$ *Marginal effect of* Age^2. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.9: Marginal Effects for Trivariate Probit Model with Cluster Corrected Standard Errors: Effects of Mortgage Size

	Government-run Lottery			Small Prize Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Mortgage quarter 2	0.04859	0.01098	<0.001	0.00154	0.00682	0.821	0.00920	0.01255	0.464
Mortgage quarter 3	0.05366	0.00936	<0.001	0.02455	0.00806	0.002	0.03695	0.01036	<0.001
Mortgage quarter 4	0.06242	0.01015	<0.001	0.01066	0.00928	0.251	0.03999	0.01162	0.001
Single male	0.01376	0.01580	0.384	-0.01800	0.01366	0.188	-0.00926	0.01485	0.533
Couple only	0.07487	0.01338	<0.001	-0.00585	0.00994	0.556	0.10915	0.01493	<0.001
Couple w/ children	0.03727	0.01548	0.016	-0.09037	0.01186	<0.001	0.18193	0.01545	<0.001
Number of vehicle	0.03866	0.00631	<0.001	0.02036	0.00341	<0.001	0.02457	0.00461	<0.001
Pri. health insurance	0.03397	0.00635	<0.001	0.00906	0.00620	0.144	0.08117	0.00790	<0.001
Life insurance	0.07229	0.00040	<0.001	0.01716	0.00019	<0.001			
Working full time	0.02653	0.00815	0.001	0.01072	0.00629	0.089	0.02227	0.00819	0.007
Metropolitan	-0.01797	0.00951	0.059	-0.00057	0.00692	0.934	-0.01909	0.00850	0.025
Age	0.02258	0.00299	<0.001	-0.01028	0.00186	<0.001	0.01755	0.00288	<0.001
Age^2	-0.00021	0.00004	<0.001	0.00009	0.00002	<0.001	-0.00015	0.00003	<0.001
Income	0.00205	0.00140	0.143	0.00773	0.00111	<0.001	0.00809	0.00168	<0.001
$Income^2$	-0.00002	0.00002	0.274	-0.00005	0.00002	0.008	-0.00001	0.00000	<0.001
HLE2	0.02629	0.01814	0.147	0.04371	0.01746	0.012	0.02670	0.01655	0.107
HLE3	0.03552	0.01900	0.062	0.03020	0.01561	0.053	0.07276	0.01775	<0.001
HLE4	-0.02641	0.01958	0.177	-0.00862	0.01829	0.637	0.10487	0.01857	<0.001
HLE5	-0.13332	0.01918	<0.001	-0.05434	0.01765	0.002	0.07976	0.01934	<0.001
Observations	17321								

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation $\textit{Marginal effect of age} + 2\overline{Age}\ \textit{Marginal effect of } Age^2$. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.10: Marginal Effects for Bivariate Probit Model with Cluster Corrected Standard Errors: Single Female Household

	Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Homeowner w/ mtg.	0.07785	0.02403	0.001	0.00968	0.02307	0.675
Number of vehicle	-0.01487	0.03158	0.638	0.02375	0.03592	0.508
Pri. health insurance	0.05350	0.04784	0.263	0.07415	0.03038	0.015
Life insurance	0.09988	0.48762	0.838			
Working full time	-0.02612	0.04244	0.538	0.03764	0.03160	0.234
Metropolitan	-0.02959	0.03672	0.420	-0.06220	0.02326	0.007
Age	0.02161	0.01554	0.165	0.01024	0.01095	0.350
Age^2	-0.00023	0.00016	0.151	-0.00009	0.00012	0.466
Income	0.01399	0.01546	0.366	0.02070	0.01336	0.121
$Income^2$	-0.00049	0.00065	0.453	-0.00075	0.00071	0.292
HLE2	-0.01954	0.05001	0.696	0.03380	0.05861	0.564
HLE3	-0.02057	0.05102	0.687	0.01809	0.05230	0.729
HLE4	-0.09845	0.07949	0.216	0.10189	0.06383	0.110
HLE5	-0.19071	0.06130	0.002	0.03599	0.06910	0.602
Observations	1569					

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation *Marginal effect of age* + $2\overline{Age}$ *Marginal effect of* Age^2. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.11: Marginal Effects for Bivariate Probit Model with Cluster Corrected Standard Errors: Single Male Household

	Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Homeowner w/ mtg.	0.06127	0.04837	0.205	0.06814	0.03004	0.023
Number of vehicle	0.01401	0.01412	0.321	0.00615	0.01654	0.710
Pri. health insurance	0.06352	0.05634	0.260	0.09677	0.02252	$<$0.001
Life insurance	0.09995	0.46934	0.831			
Working full time	0.01451	0.05735	0.800	0.10303	0.02821	$<$0.001
Metropolitan	-0.04657	0.02594	0.073	0.01975	0.02292	0.389
Age	0.01805	0.00822	0.028	0.00194	0.00960	0.840
Age^2	-0.00015	0.00009	0.089	0.00003	0.00011	0.753
Income	0.00397	0.01334	0.766	0.02723	0.00826	0.001
Income^2	-0.00005	0.00020	0.798	-0.00049	0.00025	0.053
HLE2	0.05410	0.04079	0.185	-0.03865	0.03210	0.229
HLE3	0.03329	0.04627	0.472	-0.00750	0.02912	0.797
HLE4	-0.00094	0.05753	0.987	-0.01475	0.03488	0.672
HLE5	-0.12627	0.07730	0.102	-0.08029	0.05462	0.142
Observations	1728					

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation *Marginal effect of age* + $2\overline{Age}$ *Marginal effect of Age*^{2}. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.12: Marginal Effects for Bivariate Probit Model with Cluster Corrected Standard Errors: Couple-Only Household

	Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Homeowner w/ mtg.	0.04514	0.01169	<0.001	0.00913	0.01477	0.537
Number of vehicle	0.03674	0.00959	<0.001	0.02223	0.00791	0.005
Pri. health insurance	0.03545	0.01206	0.003	0.08960	0.01079	<0.001
Life insurance	-0.03844	0.10087	0.703			
Working full time	0.01758	0.01223	0.151	0.02489	0.01266	0.049
Metropolitan	-0.00687	0.01203	0.568	-0.01745	0.01267	0.168
Age	0.01804	0.00439	<0.001	0.02025	0.00435	<0.001
Age^2	-0.00018	0.00005	<0.001	-0.00017	0.00005	0.001
Income	0.00435	0.00194	0.025	0.00787	0.00233	0.001
$Income^2$	-0.00002	0.00002	0.344	-0.00001	0.00000	0.002
HLE2	0.01001	0.02462	0.684	0.03186	0.02469	0.197
HLE3	0.02060	0.02344	0.380	0.08093	0.02518	0.001
HLE4	-0.05483	0.02609	0.036	0.06981	0.03007	0.020
HLE5	-0.14582	0.02862	<0.001	0.05890	0.03371	0.081
Observations	6670					

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation *Marginal effect of age* + $2\overline{Age}$ *Marginal effect of* Age^2. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.13: Marginal Effects for Bivariate Probit Model with Cluster Corrected Standard Errors: Couple-with-Children Household

	Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Homeowner w/ mtg.	0.05380	0.02063	0.009	0.04320	0.01827	0.018
Number of vehicle	0.05066	0.01360	<0.001	0.02679	0.00752	<0.001
Pri. health insurance	0.02113	0.02054	0.304	0.06800	0.01216	<0.001
Life insurance	0.16483	0.22030	0.454			
Working full time	0.03497	0.01054	0.001	-0.00186	0.01144	0.871
Metropolitan	-0.01465	0.01600	0.360	-0.01551	0.01279	0.225
Age	0.01022	0.01188	0.390	0.03124	0.00686	<0.001
Age^2	-0.00006	0.00014	0.640	-0.00035	0.00009	<0.001
Income	0.00278	0.00457	0.543	0.01641	0.00225	<0.001
$Income^2$	-0.00003	0.00004	0.510	-0.00012	0.00004	0.001
HLE2	0.05476	0.04737	0.248	0.08115	0.04273	0.058
HLE3	0.06824	0.05740	0.235	0.14169	0.03972	<0.001
HLE4	0.00604	0.06175	0.922	0.19057	0.03867	<0.001
HLE5	-0.10676	0.05233	0.041	0.16437	0.04075	<0.001
Observations	7353					

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation *Marginal effect of age* + $2\overline{Age}$ *Marginal effect of* Age^2. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.14: Marginal Effects for Trivariate Probit Model with Cluster Corrected Standard Errors: Single Female Household

	Government-run Lottery			Small Prize Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Mortgage quarter 2	0.05364	0.02987	0.073	0.05513	0.02350	0.019	0.04072	0.02725	0.135
Mortgage quarter 3	0.11912	0.03052	<0.001	0.00886	0.02492	0.722	0.00511	0.02854	0.858
Mortgage quarter 4	0.03997	0.04025	0.321	0.03906	0.03396	0.250	-0.07868	0.04481	0.079
Number of vehicle	-0.00257	0.02903	0.929	-0.02282	0.02322	0.326	0.02450	0.03499	0.484
Pri. health insurance	0.05056	0.02492	0.042	0.02285	0.01757	0.193	0.07291	0.02999	0.015
Life insurance	0.06488	0.00058	<0.001	0.02876	0.00088	<0.001			
Working full time	-0.01714	0.03610	0.635	-0.03769	0.02359	0.110	0.03241	0.03142	0.302
Metropolitan	-0.02084	0.03339	0.533	-0.02610	0.02208	0.237	-0.05486	0.02410	0.023
Age	0.03470	0.01201	0.004	-0.00705	0.00846	0.405	0.01012	0.01108	0.361
Age^2	-0.00035	0.00013	0.006	0.00005	0.00009	0.598	-0.00009	0.00012	0.469
Income	0.00330	0.01216	0.786	0.03620	0.01058	0.001	0.02655	0.01358	0.051
Income^2	-0.00003	0.00056	0.958	-0.00170	0.00063	0.007	-0.00089	0.00071	0.211
HLE2	-0.04199	0.04722	0.374	0.05742	0.04493	0.201	0.03188	0.05568	0.567
HLE3	-0.03573	0.04850	0.461	0.02697	0.04775	0.572	0.01500	0.05007	0.765
HLE4	-0.10329	0.06254	0.099	-0.02465	0.04466	0.581	0.10340	0.06143	0.092
HLE5	-0.20172	0.06117	0.001	-0.08029	0.05554	0.148	0.03773	0.06718	0.574
Observations	1569								

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation $\textit{Marginal effect of age} + 2\overline{Age}\ \textit{Marginal effect of } Age^2$. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.15: Marginal Effects for Trivariate Probit Model with Cluster Corrected Standard Errors: Single Male Household

	Government-run Lottery			Small Prize Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Mortgage quarter 2	0.07028	0.02830	0.013	-0.02203	0.02810	0.433	0.04490	0.03488	0.198
Mortgage quarter 3	0.06810	0.02790	0.015	-0.02186	0.02380	0.358	0.06853	0.03517	0.051
Mortgage quarter 4	0.09182	0.04152	0.027	-0.01213	0.03475	0.727	0.05836	0.04102	0.155
Number of vehicle	0.00894	0.01454	0.539	0.00028	0.01056	0.979	0.00611	0.01639	0.709
Pri. Health insurance	0.06074	0.02233	0.007	0.02152	0.02293	0.348	0.09747	0.02177	<0.001
Life insurance	0.07541	0.00090	<0.001	-0.01429	0.00048	<0.001			
Working full time	0.02929	0.03028	0.333	-0.01163	0.02268	0.608	0.10405	0.02776	<0.001
Metropolitan	-0.07299	0.02663	0.006	-0.00107	0.02358	0.964	0.01843	0.02243	0.411
Age	0.01986	0.00775	0.010	-0.01736	0.00674	0.010	0.00203	0.00961	0.833
Age^2	-0.00014	0.00009	0.107	0.00018	0.00008	0.023	0.00003	0.00011	0.777
Income	0.00042	0.00615	0.946	0.01590	0.00436	<0.001	0.02706	0.00682	<0.001
$Income^2$	-0.00004	0.00012	0.754	-0.00021	0.00008	0.012	-0.00049	0.00020	0.015
HLE2	0.09002	0.03666	0.014	0.03154	0.03339	0.345	-0.03997	0.03227	0.215
HLE3	0.06038	0.04292	0.159	0.00235	0.02974	0.937	-0.00817	0.02947	0.782
HLE4	0.01893	0.05551	0.733	-0.01716	0.03740	0.646	-0.01696	0.03467	0.625
HLE5	-0.14557	0.05310	0.006	-0.00040	0.04224	0.993	-0.08255	0.05524	0.135
Observations	1728								

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation $\textit{Marginal effect of age} + 2\overline{Age}\ \textit{Marginal effect of } Age^2$. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.16: Marginal Effects for Trivariate Probit Model with Cluster Corrected Standard Errors: Couple-Only Household

	Government-run Lottery			Small Prize Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Mortgage quarter 2	0.03833	0.01677	0.022	-0.00750	0.01423	0.598	-0.00417	0.01916	0.828
Mortgage quarter 3	0.04396	0.01509	0.004	0.03448	0.01276	0.007	0.00438	0.01773	0.805
Mortgage quarter 4	0.05455	0.01544	$<$0.001	0.00817	0.01548	0.598	0.02665	0.01903	0.161
Number of Vehicle	0.03804	0.00920	$<$0.001	0.01513	0.00676	0.025	0.02227	0.00788	0.005
Pri. Health Insurance	0.02730	0.01176	0.020	0.00808	0.00998	0.418	0.08937	0.01077	$<$0.001
Life Insurance	0.07906	0.00083	$<$0.001	0.01273	0.00019	$<$0.001			
Working full time	0.01471	0.01257	0.242	0.01697	0.00995	0.088	0.02338	0.01256	0.063
Metropolitan	-0.00637	0.01207	0.598	0.00993	0.01092	0.363	-0.01972	0.01285	0.125
Age	0.02148	0.00440	$<$0.001	-0.01107	0.00362	0.002	0.02020	0.00434	$<$0.001
Age^2	-0.00021	0.00005	$<$0.001	0.00010	0.00004	0.020	-0.00017	0.00005	0.001
Income	0.00154	0.00179	0.388	0.00718	0.00148	$<$0.001	0.00759	0.00231	0.001
$Income^2$	-0.00001	0.00002	0.714	-0.00004	0.00002	0.057	-0.00001	0.00000	0.003
HLE2	0.00627	0.02602	0.810	0.04483	0.03003	0.136	0.03243	0.02457	0.187
HLE3	0.01896	0.02237	0.397	0.02810	0.02582	0.277	0.08141	0.02523	0.001
HLE4	-0.05035	0.02445	0.039	-0.02213	0.03158	0.483	0.06837	0.03021	0.024
HLE5	-0.14216	0.02630	$<$0.001	-0.06487	0.03113	0.037	0.05763	0.03393	0.089
Observations	6670								

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation $\textit{Marginal effect of age} + 2\overline{Age}\ \textit{Marginal effect of Age}^2$. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.17: Marginal Effects for Trivariate Probit Model with Cluster Corrected Standard Errors: Couple-with-Children Household

	Government-run Lottery			Small Prize Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Mortgage quarter 2	0.05475	0.02328	0.019	-0.00166	0.01376	0.904	0.01410	0.02132	0.508
Mortgage quarter 3	0.04901	0.01619	0.002	0.02396	0.01093	0.028	0.05920	0.02025	0.003
Mortgage quarter 4	0.07008	0.01621	$<$0.001	0.01139	0.01163	0.327	0.04657	0.01993	0.019
Number of vehicle	0.04839	0.00957	$<$0.001	0.03193	0.00650	$<$0.001	0.02666	0.00748	$<$0.001
Pri. health insurance	0.03131	0.01025	0.002	0.00404	0.00855	0.636	0.06841	0.01206	$<$0.001
Life insurance	0.06549	0.00046	$<$0.001	0.02588	0.00051	$<$0.001			
Working full time	0.03694	0.01079	0.001	0.01132	0.00805	0.160	-0.00335	0.01144	0.770
Metropolitan	-0.01712	0.01440	0.234	-0.00448	0.00879	0.611	-0.01950	0.01281	0.128
Age	0.01703	0.00686	0.013	-0.01084	0.00491	0.027	0.03039	0.00694	$<$0.001
Age^2	-0.00014	0.00009	0.126	0.00011	0.00006	0.090	-0.00034	0.00009	$<$0.001
Income	0.00233	0.00201	0.248	0.00704	0.00154	$<$0.001	0.01599	0.00233	$<$0.001
$Income^2$	-0.00003	0.00003	0.368	-0.00004	0.00002	0.076	-0.00012	0.00004	0.001
HLE2	0.06317	0.04272	0.139	0.07203	0.03048	0.018	0.08415	0.04251	0.048
HLE3	0.08094	0.04216	0.055	0.06767	0.02839	0.017	0.14236	0.03924	$<$0.001
HLE4	0.02428	0.04344	0.576	0.03338	0.02827	0.238	0.18949	0.03839	$<$0.001
HLE5	-0.09184	0.04351	0.035	-0.02209	0.03082	0.473	0.16416	0.04041	$<$0.001
Observations	7353								

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation $\textit{Marginal effect of age} + 2\overline{Age}\ \textit{Marginal effect of } Age^2$. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.18: Marginal Effects for Bivariate Probit Model with Cluster Corrected Standard Errors: Small Spending Adjusted

	Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Homeowner w/ mtg.	0.05459	0.00864	<0.001	0.03003	0.00940	0.001
Single male	0.00829	0.01484	0.576	-0.00836	0.01493	0.575
Couple only	0.06884	0.02138	0.001	0.11206	0.01513	<0.001
Couple w/ children	0.01929	0.03049	0.527	0.18583	0.01555	<0.001
Number of vehicle	0.04167	0.00715	<0.001	0.02481	0.00460	<0.001
Pri. health insurance	0.03722	0.01231	0.002	0.08118	0.00789	<0.001
Life insurance	0.01617	0.14531	0.911			
Working full time	0.02797	0.00889	0.002	0.02352	0.00823	0.004
Metropolitan	-0.01526	0.01013	0.132	-0.01622	0.00845	0.055
Age	0.01964	0.00392	<0.001	0.01767	0.00288	<0.001
Age^2	-0.00019	0.00004	<0.001	-0.00015	0.00003	<0.001
Income	0.00514	0.00189	0.007	0.00836	0.00168	<0.001
$Income^2$	-0.00004	0.00002	0.033	-0.00001	0.00000	<0.001
HLE2	0.02158	0.01829	0.238	0.02672	0.01665	0.109
HLE3	0.03280	0.02203	0.136	0.07354	0.01789	<0.001
HLE4	-0.03357	0.02511	0.181	0.10751	0.01854	<0.001
HLE5	-0.14358	0.02395	<0.001	0.08255	0.01943	<0.001
Observations	17321					

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation *Marginal effect of age* + $2\overline{Age}$ *Marginal effect of* Age^2. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Table 2.19: Marginal Effects for Trivariate Probit Model with Cluster Corrected Standard Errors: Small Spending Adjusted

	Government-run Lottery			Small Prize Gambling			Life Insurance		
	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value	M.E.	Std. Err.	p-value
Mortgage quarter 2	0.04755	0.01077	<0.001	0.00336	0.00680	0.621	0.00920	0.01255	0.464
Mortgage quarter 3	0.05425	0.00962	<0.001	0.02458	0.00803	0.002	0.03695	0.01036	<0.001
Mortgage quarter 4	0.05917	0.01057	<0.001	0.01163	0.00942	0.217	0.03999	0.01162	0.001
Single male	0.01272	0.01662	0.444	-0.01651	0.01355	0.223	-0.00926	0.01485	0.533
Couple only	0.07712	0.01341	<0.001	-0.00631	0.00995	0.526	0.10915	0.01493	<0.001
Couple w/ children	0.04034	0.01555	0.009	-0.08874	0.01162	<0.001	0.18193	0.01545	<0.001
Number of vehicle	0.03880	0.00618	<0.001	0.01966	0.00332	<0.001	0.02457	0.00461	<0.001
Pri. health insurance	0.03393	0.00651	<0.001	0.00789	0.00609	0.195	0.08117	0.00790	<0.001
Life insurance	0.07105	0.00039	<0.001	0.01739	0.00019	<0.001			
Working full time	0.02794	0.00834	0.001	0.00976	0.00633	0.123	0.02227	0.00819	0.007
Metropolitan	-0.01833	0.00969	0.059	0.00085	0.00693	0.902	-0.01909	0.00850	0.025
Age	0.02254	0.00297	<0.001	-0.01017	0.00187	<0.001	0.01755	0.00288	<0.001
Age^2	-0.00021	0.00003	<0.001	0.00009	0.00002	<0.001	-0.00015	0.00003	<0.001
Income	0.00262	0.00140	0.061	0.00769	0.00109	<0.001	0.00809	0.00168	<0.001
$Income^2$	-0.00003	0.00002	0.205	-0.00005	0.00002	0.008	-0.00001	0.00000	<0.001
HLE2	0.02567	0.01818	0.158	0.04154	0.01737	0.017	0.02670	0.01655	0.107
HLE3	0.03429	0.01891	0.070	0.02847	0.01546	0.066	0.07276	0.01775	<0.001
HLE4	-0.03161	0.01947	0.105	-0.00993	0.01820	0.585	0.10487	0.01857	<0.001
HLE5	-0.14247	0.01922	<0.001	-0.05732	0.01740	0.001	0.07976	0.01934	<0.001
Observations	17321								

Note:

1. Marginal effects of *Age* should take into account effects from both Age and Age^2. Marginal effects for average age can be evaluated approximately using equation *Marginal effect of age* + $2\overline{Age}$ *Marginal effect of* Age^2. $\overline{Age}$ is the average of age. Marginal effects for average income can be evaluated in much the same way.

Chapter 3

The Relationship Between Consumer Spending on Exercise, Sports Betting and Attending Sporting Events

3.1 Introduction

Recent developments in the United States and Canada highlight the importance of understanding the relationship among consumer spending on leisure activities like spectator sporting events, gambling, and participation in physical activity. Rates of participation in physical activity have dropped continuously in recent decades in developed economies, suggesting that consumers devote more time and resources to other leisure time activities. At the same time, access to legal gambling has increased significantly, both at casinos and on-line. Sports betting markets in North America are in a state of transition, with important increases and decreases in sports betting opportunities occurring frequently. In Canada, the federal government has debated an amendment to the Criminal Code that would legalize betting on individual sporting events. In the US, the State of New Jersey has attempted to legalize sports betting; in response, all major North American professional sports leagues and the NCAA sued the state, citing economic damages to their revenues among the possible reasons for blocking this expansion of legal sports betting. These events suggest important links between consumer spending on spectator sporting events, participation in leisure time physical activity, and gambling. This paper focuses on the relationship between spending on sports betting and other related categories of consumer spending in order to assess how increasing access to legal sports betting opportunities can directly affect household spending and indirectly affect other sectors of the economy.

We investigate the relationship between consumer spending on spectator sporting events, leisure-time physical activity and sports betting using data from the Survey of Household Spending (SHS), a large scale survey of consumer economic activity conducted annually in Canada. We use these data to estimate the parameters of several versions of an Almost Ideal Demand System (AIDS) for these consumer goods and services. AIDS models, or the related Quadratic AIDS (QAIDS) model have been used extensively to analyze patterns in consumer spending. Recent applications include gasoline and consumer transportation spending (Chang and Serletis, 2013), consumer travel (Mangion et al., 2012), consumer non-durables (Blow et al., 2012), food (Koohi-Kamali, 2013), and residential energy (Guta, 2012; Gebreegziabher et al., 2012).

Flexible functional form demand models like AIDS and QAIDS models allow for the estimation of own and cross price elasticities for consumer goods and services that can be interpreted as approximations to general consumer choice models. These models generate estimates consistent with aggregation across households, which makes them ideal for use with consumer expenditure survey data like the SHS data used here.

Our results suggest that household spending on lottery based sports betting and spectator sports are complements, but household spending on lottery based sports betting and spending on leisure time physical activity and exercise are substitutes, based on household spending data from more than 145,000 Canadian households that participated in the SHS over the period 1997-2009. Since household spending on sports betting and spending on attendance at sporting events are complements, professional sports leagues in North America have little reason to worry that the expansion of legal sports betting opportunities will lead to reductions in revenues generated by games; instead, the evidence developed here suggests that expanding opportunities for legal sports betting may increase attendance at live sporting events, generating larger revenues for sports teams and leagues. However, the evidence that consumer spending on sports betting and spending on exercise and physical activity suggests that increasing access to legal sports betting may have unintended consequences, if the reduced household spending on physical activity and exercise reflects reduced participation in physical activity, since reduced participation in physical activity has been linked to adverse health outcomes (Humphreys et al., 2014).

3.2 Betting, Live Game Attendance, and Consumer Spending

Gambling, including sports betting, is a highly regulated economic activity. The other two categories of consumer spending analyzed here, spending on live attendance at sporting events, and spending on leisure time physical activity, are not regulated; professional sports, and Division I intercollegiate athletics in the US, are produced by monopoly sports leagues, and by a cartel in the case of NCAA sports. Access to sports betting depends on both government regulations and the willingness of certain individuals to violate these regulations. Simmons (2007) examined the factors that influence gambling regulations. Simmons (2007) identified the fundamental tension between consumers, who view gambling as entertainment and a financial transaction, governments, who view state sponsored monopoly gambling as an important revenue source, and, in the case of sports betting, sports teams and leagues, as a key factor affecting access to legal gambling. In many countries with limited access to legal sports betting opportunities, significant illegal sports betting operations exist. Strumpf (2003) developed evidence of extensive illegal sports betting operations in the US.

Sauer (2001) analyzed the regulation and availability of legal gambling using a public choice model in which governments set regulations in response to lobbying by interest groups, including pro-gambling consumers whose welfare rises with gambling access and falls with gambling regulation and an anti-gambling lobbying group who want to restrict gambling access. The anti-gambling lobbying group contains individuals and organizations like churches that disapprove of gambling for differing reasons. In the case of sports betting, the anti-gambling group can also contain professional and amateur sports organizations like the NFL, NBA, NHL and NCAA. The gambling regulations predicted by this model depend on the relative lobbying efforts made by the two groups.

Forrest and Simmons (2003) reviewed the economic and public policy context for sports betting. They highlighted the recent and rapid growth in sports betting world wide and discussed the implications for revenue generation for government and sports organizations. Forrest and Simmons (2003) identified a number of negative aspects of sports betting, including incentives for corruption. Forrest and Simmons (2003) emphasized the symbiotic nature of the relationship between sport and sports betting and pointed out the importance of complementarities between watching sport and sports betting as well as the tensions generated by this relationship. The importance of complementarities in consumption drives demand for sports betting and puts pressure on governments to expand sports betting opportunities while the corruptive factors fuel the desires of anti-sports gambling groups and leads to increased pressure to restrict sports betting opportunities.

Even though gambling, and sports betting, is highly regulated, in both Canada and the United States, government regulators are currently debating policy changes that would significantly increase access to legal sports betting in both countries. In Canada, the Criminal Code currently prohibits single-event sports betting but allows parlay-style betting on two or more sporting events. However, Bill C-290 ("an Act to amend the Criminal Code") which would legalize single-event sports betting was passed unanimously by the Canadian House of Commmons in December 2012 and has been passed to, but not taken up, by the Senate. The governments of eight Canadian provinces have formally supported passage of the bill. The National Hockey League has vocally opposed the bill, claiming that legal single-event sports betting will lead to reduced revenues and increase the likelihood that gamblers will try to fix game outcomes. During legislative hearings on the bill, experts testified that Canadians currently spend between \$10 and \$40 billion annually betting on single sporting events with illegal bookmakers and about \$4 billion annually betting on single sporting events with on-line off-shore bookmakers that are not regulated by Canadian provinces.

In the US, the State of New Jersey attempted to pass a law legalizing single-event sports betting in 2012 following a successful 2011 referendum on the issue. Single-event sports betting is currently legal at casinos in Las Vegas, and under the federal Professional and Amateur Sports Protection Act (PASPA) of 1993 sports betting can be legalized in four states (Nevada, Oregon, Delaware and Montana). The four major professional sports leagues operating in the US, the National Football League (NFL), Major League Baseball (MLB), the National Basketball Association (NBA) and the National Hockey League (NHL) along with the National Collegiate Athletic Association (NCAA) challenged New Jersey's legalization of sports betting immediately after the law was passed and obtained an injunction against the implementation of the law and appear to have blocked its implementation. The sports leagues claimed that increased access to sports betting would damage the public perception of professional sports, leading to long-term and irreparable economic damage in the form of lost revenues.

Little evidence about the relationship between legal sports betting and professional sports leagues currently exists. Forrest and Pérez (2011) show that the presence of high-profile football matches on the betting coupon increases the volume of bets placed in *La Quiniela*, a football pool betting game operated buy the Spanish National Lottery. García and Rodríguez (2007) and García et al. (2008) reported evidence of important complementarities between watching or following sporting events and sports betting in *La Quiniela* in Spain. No formal evidence of such complementarities exists in other countries.

Betting on single-event sports is currently legal in the UK, much of the European Union, and Australia. In the US, single-event sports betting is legal only at casinos in Nevada.

Canadians have access to legal sports betting as part of a group of government-sanctioned lottery games. These games have different names in different parts of Canada. In Quebec, the lottery based sports betting game is called Pari Sportif; in Ontario and Atlantic Canada this game is called Pro-Line; in Manitoba, Saskatchewan and Alberta it is called Sports Select; in British Columbia it is called Sports Action. In western Canada, consumers also have access to a lottery based game based on point spreads. All of these lottery based sports betting games are available at lottery outlets. In some provinces, sports betting and other lottery tickets can be purchased on the internet. In 2011, total sales of government sponsored lottery products in Canada was $4.76 billion dollars. Canadians have easy access to lottery outlets. About $237 million of these sales came from sports betting games, just over 5% of total sales. In 2011 there were 30,090 lottery outlets in Canada, roughly one lottery outlet for every 900 persons age 18 and older.

These Canadian lottery based sports betting games, except Point Spread, are based on fixed odds bets on game outcomes and total points scored in professional and amateur sports leagues, including games in the major North American sports leagues and US college football and basketball games. These games feature parlay-style betting where gamblers must pick the outcome of between two and twelve games.

Winnings in Canadian sports betting games are not pari-mutuel; lottery corporations make profits based on over-round, the amount by which the win probabilities implied by the fixed betting odds offered on specific outcomes exceed 100. The over-round on Canadian sports betting games varies with the number of events selected in the lottery ticket. The minimum over-round is 160%, and it can exceed 300% depending on the bets placed. Winnings are capped at $2,000,000 per card no matter how large the odds on the selected events or the number of games included in the parlay wager.

Given that Canadians currently have access to a form of sports betting offered at lottery outlets throughout the country, data on household spending on government operated lotteries and other related consumer goods can provide new insights into the relationship between sports betting and other sectors of the economy like professional sports. In the following sections, we describe the data source, empirical methods, and results of this analysis.

We also analyze the relationship between consumer spending on live attendance at sporting events and spending on leisure time physical activity and exercise. Professional sports leagues in North America often sponsor programs to encourage participation in physical activity. For example, the NFLs "Play 60" program attempts to encourage youth participation in physical activity. In an international setting, hosts of the Olympic Games often claim that hosting the Games generates a "demonstration effect" through which exposure to the Games, through attendance or advertizing, leads local residents to become more physically active. Analyzing patterns in consumer spending on attending live sporting events and spending on leisure time physical activity can help to assess the validity of such claims.

3.3 Data Description

The data come from the Survey of Household Spending (SHS) conducted annually by Statistics Canada. We analyze data from the SHS confidential micro data files from 1997 to 2009, which includes detailed information about household expenditure and characteristics. In particular, the SHS contains information about household expenditure on attendance at live sporting events, physical activity/physical fitness, and government-run lottery as well as detailed household demographic characteristics.

The SHS has been conducted in the ten Canadian provinces annually since 1997 and in the three Canadian territories biannually since 1999. The coverage of the population is about 98% in the provinces and 90% in the territories. However, there was a drop in coverage in Nunavut in 2005 (68.3%) due to changes in sampling methods. The SHS contained information about 20,000 eligible households in each survey year before 2007. The sample size was reduced by nearly 30% after 2008 due to Statistics Canada budget cuts. The response rates for the SHS have been relatively high, about 66% over this period, varying from 63.4% in 2008 to 76.2% in 2001.

The SHS contains detailed information about household expenditure for the reference (calendar) year is collected during a personal interview. A paper questionnaire was used before 2006 and replaced by computer assisted personal interviews since 2006. Also beginning in 2006, the SHS collected information about dwelling characteristics and household equipment at the time of the interview instead of the end of the reference year. Another important change since 2006 was that there was no distinction between "part-year" and "full-year" members and households. Data were collected for every household member as of the time of the interview. Persons temporarily living away from their families (for example, students at university) were included in the household to avoid double counting.

The SHS sample uses stratification and multi-stage selection from the Labour Force Survey (LFS) sampling frame. From 1997 to 2003 SHS used 1991 Census geography and 1991 population counts; from 2004, SHS used 2001 Census geography and 2001 population counts after the major sample redesign of the LFS. SHS also used Census of Population and T4 data from the Canada Revenue Agency to adjust the survey weights. For comparisons over years, data from the 1997 to 2003 SHS were also re-weighted using the new weighting methodology. However, since the SHS employed a complex survey design and also considered uneven respondent selection probabilities, estimation and variance calculations using the weighted sample over years for certain subgroups or variables may require extra caution.

Extra questions were included in several years, but none of these relate to the variables we are interested. Our data set contains all SHS households reporting positive total expenditure on attendance at live sporting events, physical activity and exercise, and government-run lotteries (which includes sports betting) from the SHS confidential micro data files. The sample contains 145,560 observations including data from tens of thousands of Canadian households over the years 1997 to 2009. Due to the nature of confidential data, we cannot determine whether the same households participated in the surveys in multiple years. Thus, we assume that all 145,560 observations were from different households. Summary statistics are shown on Table

3.1. Note that we can not report the maximum and minimum values due to the Statistics Canada confidentiality policy. Instead, we report the values at the 0.02% percentile and 99.98%, which can be almost regarded as the minimum and maximum values.

The reported *Real Expenditure* on Table 3.1 is the total real household expenditure for (1) admissions to live sporting events, (2) total fees (including membership and single usage fees) and dues for sports activities, sports and recreation facilities, and health clubs, and (3) government-run lotteries. It varies from about 1 dollar to over 15,000 dollars. Table 3.1 also shows the shares of total expenditures for these three goods and services. On average, the share of spectator sports event attendance is the smallest, only about 10% of this spending. The spending on government-run lotteries and sports betting is the largest, accounting for over one half of this spending. For each of the three goods and services, there were always some households that spent nothing on some of these (share = 0) or only spent money on one of the goods and services (share = 1).

Real Household Income is the household's total annual income before taxes in the reference year. The average income in our sample is just above 62 thousand; the top earning households in the sample earned over 1 million. *Reference Person Age* is the age of reference person in the reference year. For one-person households or single-parent households, the household income is mainly earned by the reference person. From an economic point of view the age of the reference person can be regarded as the age of the household. For couples, the age of the spouse of the reference person is highly correlated with the age of the reference person. As a matter of fact, the ages of couples are usually very close to each other. Hence we can use the reference person's age as the age of the household.

The rest of the variables on Table 3.1 are all dummy variables. *Urban* identifies whether the household lives in an urban area. Most households (78.6%) in our sample lived in urban areas. The definition of urban area follows the census definition: "minimum population concentrations of 1,000 and a population density of at least 400 per square kilometre." The other 21 variables are categorized in 4 groups: housing tenure, major income source, household type and household employment status.

The SHS identifies 5 types of housing tenure. The shares of households with and without a mortgage are both about 1/3. Regular renters account for a bit more than one quarter of the sample. There were also nearly 1% special tenants who don't pay rent. Roughly 3% of households were in the mixed tenure group. They both owned and rented their dwellings in the reference year. Major household income sources are divided into 6 groups. Over 2/3 of the households reported income from paid employment as the major source of income. Less than 20% of households depended on government transfer payments as the major source of income. A significant part of this subgroup were retired.

The SHS defines 4 basic household types: one-person households, couples, single-parent households and other households. We further broke households down into 7 types: single female households, single male households, couples only households, couples with children and/or other persons, single mothers, single fathers and other households. The rationale here was to separately identify the impacts of gender and children. In the 4th type, couples with children and/or other person(s) living in the household, other persons refer to children whose marital status is not "single, never-married," relatives by birth or marriage, and unrelated persons. The definition is complex, but basically most households in this type are just couples with children. There were about 20% one-person households in our data. Single females and single males constitute similar shares of the sample. Couples accounted for almost 2/3 of the sample. Couples with children households were more common than couple only households. There were 6% single mother households, about 4 times of the single father households.

Household employment status consists of 3 variables identifying 3 different labour market outcomes: (1) both the reference person and the spouse have no job (if reference person had a married or common law spouse); (2) at least one of the couple has a part time job; and (3) both of the couple have full time jobs. The largest group in the sample is type (2) households. In 46.1% of the sample households, at least one person was working during the reference year, but both members of the couple did not have full time jobs. Also a little more than 20% households had no job at all.

Table 3.1: Summary Statistics

Variable	Mean	Std. Dev.	0.02nd percentile	99.98th percentile
Budget share, admissions to live sports events	0.101	0.237	0	1
Budget share, expenditure on physical activities	0.364	0.411	0	1
Budget share, expenditure on government-run lottery	0.535	0.435	0	1
Real Expenditure	518	847	1.02	15734
Reference Person Age	48.3	15.5	17	94
Real Household Income (000)	62.41	79.38	0	1163.52
Urban Household	0.786	—	0	1
Housing Characteristics				
Homeowner without mortgage	0.334	—	0	1
Homeowner with mortgage	0.360	—	0	1
Tenants - regular	0.266	—	0	1
Tenants - rent-free	0.009	—	0	1
Mixed type	0.031	—	0	1
Major Income Source				
No income	0.001	—	0	1
Paid employment	0.664	—	0	1
Self-employment	0.065	—	0	1
Investment income	0.014	—	0	1
Government transfer payments	0.185	—	0	1
Miscellaneous	0.071	—	0	1
Household Type				
Single Female	0.110	—	0	1
Single male	0.097	—	0	1
Couple only	0.276	—	0	1
Couple with children	0.379	—	0	1
Single mother	0.060	—	0	1
Single father	0.014	—	0	1
Other	0.063	—	0	1
Household Employment Status				
Reference person and spouse not working	0.213	—	0	1
At least one employed part time	0.461	—	0	1
Both employed full-time	0.325	—	0	1

3.4 Empirical Approach

Deaton and Muellbauer (1980) proposed the Almost Ideal Demand System (*AIDS*) model to analyze consumer spending decisions. An AIDS model represents "an arbitrary first-order approximation to any demand system" [Deaton and Muellbauer (1980)]. As Deaton and Muellbauer (1980) point out, the AIDS model satisfies all standard axioms of choice and also aggregate perfectly over consumers. The AIDS model has a simple functional form that makes it easy to estimate.

The budget share form of an AIDS demand function is the primary functional form used in the literature. This demand function is

$$w_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log \left(\frac{m}{a(p)} \right) \tag{3.1}$$

where w_i is expenditure on good or service i, the p_is are prices, m is total expenditure, and α_i, γ_{ij} and β_i are vectors of parameters. The price index $a(p)$ is defined by

$$\log a(p) = \alpha_0 + \sum_k \alpha_k \log p_k + \frac{1}{2} \sum_j \sum_k \gamma_{kj} \log p_k \log p_j \tag{3.2}$$

where again p_j is the price of good j[1]. There are three restrictions on the parameters of Equation (3.1), which ensure the demand function system satisfies: (1) the sum of the share functions equals 1 ($\sum w_i = 1$); (2) each share function is homogeneous of degree zero in prices and (3) Slutsky symmetry. Formally

$$\sum_{i=1}^{n} \alpha_i = 1, \sum_{i=1}^{n} \beta_i = 0, \sum_{i=1}^{n} \gamma_i = 0 \tag{3.3}$$

$$\sum_j \gamma_{ij} = 0 \tag{3.4}$$

$$\gamma_{ij} = \gamma_{ji} \tag{3.5}$$

[1] In our analysis, we used the component of the Consumer Price Index for spectator entertainment (excluding cablevision and satellite services) as the price of attending a spectator sports event, and the price index component for use of recreational facilities and services as the price of physical activity. Since the price of the government-sponsored sports lottery was constant over the period, 1 or 2 dollars per ticket, we used the inverse of the CPI as an approximation to this price. Unfortunately, we lack sufficient data on odds-overround to calculate an effective price index for lottery ticket spending.

The effect of changes in prices on budget shares is captured by the vector of parameters γ_{ij} and changes in total expenditure (m) are captured by the parameter vector β_i. A positive β_i identifies luxury goods and a negative β_i identifies necessities.

Banks et al. (1997) argued that, for certain goods, a log-linear expenditure share model does not provide an accurate description of consumer behavior. They derived a demand system that included a quadratic term in log expenditure, called the Quadratic Almost Ideal Demand System (QAIDS). This model is constructed in a way to nest the AIDS model and has log-linear expenditure as the left-hand side variable.

The QAIDS model comes from the indirect utility function

$$\ln V = \left(\left[\frac{\ln m - \ln a(p)}{b(p)} \right]^{-1} + \lambda(p) \right)^{-1} \tag{3.6}$$

where $a(p)$ takes the same translog form as in the AIDS model, $b(p)$ is the Cobb-Douglas price aggregator, defined by

$$b(p) = \prod_{i=1}^{n} p_i^{\beta_i} \tag{3.7}$$

and $\lambda(p)$ is simply

$$\lambda(p) = \sum_{i=1}^{n} \lambda_i \ln p_i. \tag{3.8}$$

the sum of the budget shares, $\sum w_i = 1$. The QAIDS model requires a different adding-up constraint from the AIDS model

$$\sum_i \lambda_i = 0 \tag{3.9}$$

and the corresponding system of budget share equations for the QAIDS model is

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln \left[\frac{m}{a(p)} \right] + \frac{\lambda_i}{b(p)} \left(\ln \left[\frac{m}{a(p)} \right] \right)^2 \tag{3.10}$$

Banks et al. (1997) showed that the original AIDS model is a special case of the QAIDS model when $\lambda_i = 0$ for all i. Therefore, choosing between the AIDS and the QAIDS can be simply done by testing the null hypothesis $H_0 : \lambda_i = 0 \forall i$.

Deaton and Muellbauer (1980) observe that if individual prices are highly collinear, the variation in the log price index $\ln[a(p)]$ would be reduced making the identification of α_0 difficult. They suggested assigning an *a priori* value to α_0. Banks et al. (1997) further argued that the lowest value of log expenditure places an upper bound on α_0. Here, we assign the parameter α_0 the largest integer just below the minimum level of $\ln m$ in the sample. We also tested a grid of values for α_0 and all our results remained robust to this change.

Banks et al. (1997) developed an approach that allowed demographic and household characteristics to enter the QAIDS model, Equations (3.10), through α_i. However, they also argued these characteristics could be incorporated in all terms using price scaling techniques since the AIDS/QAIDS demand systems are invariant to rescaling. In this paper, we employed the price scaling technique introduced by Ray (1983), who incorporated demographic characteristics into the original AIDS model in his research on the cost of children.

The expenditure of a representative household R is denoted by $e^R(p,u)$. u is the utility level at a given price p and z is a vector of demographic characteristics. Ray (1983) used a function $m_0(p,z,u)$ to scale the expenditure function $e^R(p,u)$ which incorporates the demographic characteristics.

$$e(p,z,u) = m_0(p,z,u) \times e^R(p,u) \tag{3.11}$$

The scaling function can be decomposed into two factors:

$$m_0(p,z,u) = \overline{m_0}(z) \times \phi(p,z,u) \tag{3.12}$$

The first component $\overline{m_0}$ only controls for changes in household characteristics, ignoring changes in consumption patterns. Ray (1983) defined the functional form for $\overline{m_0}$ as:

$$\overline{m_0}(z) = 1 + \rho' z \tag{3.13}$$

where ρ is a vector of parameters indicating the impact of household characteristics at base year ($p = 1$).

For the second term, we adopted the functional form as parameterized by Poi (2012):

$$\ln \phi(p, z, u) = \frac{\prod_{j=1}^{k} p_j^{\beta_j} (\prod_{j=1}^{k} p_j^{\eta_j' z} - 1)}{\frac{1}{u} - \sum_{j=1}^{k} \lambda_j \ln p_j}. \tag{3.14}$$

Therefore, the system of budget share equations for the QAIDS model with demographic characteristics is

$$w_i = \alpha_i + \sum_{j=1}^{k} \gamma_{ij} \ln p_j + (\beta_i + \eta_i' z) \ln \left(\frac{m}{\overline{m_0}(z) a(p)} \right) + \frac{\lambda_i}{b(p) c(p, z)} \left[\ln \left(\frac{m}{\overline{m_0}(z) a(p)} \right) \right]^2 \tag{3.15}$$

where

$$c(p, z) = \prod_{j=1}^{k} p_j^{\eta_j' z} \tag{3.16}$$

and the new adding-up constraint becomes

$$\sum_{j=1}^{k} \eta_j = 0 \tag{3.17}$$

When $\lambda_i = 0$, this model collapses to the standard AIDS model augmented with demographics developed by Ray (1983). Differentiating Equation (3.15) with respect to $\ln m$ generates the expenditure elasticity for good i

$$\mu_i = 1 + \frac{1}{w_i} \left[\beta_i + \eta_i' z + \frac{2\lambda_i}{b(p) c(p, z)} \ln \left(\frac{m}{\overline{m_0}(z) a(p)} \right) \right]. \tag{3.18}$$

Differentiating Equation (3.15) with respect to $\ln p_j$ generates the uncompensated price elasticity of good i with respect to changes in the price of good j

$$\begin{aligned} \epsilon_{ij} = -\delta_{ij} + \frac{1}{w_i} \Bigg(\gamma_{ij} - \left[\beta_i + \eta_i' z + \frac{2\lambda_i}{b(p) c(p, z)} \ln \left(\frac{m}{\overline{m_0}(z) a(p)} \right) \right] \\ \times (\alpha_j + \sum_l \gamma_{jl} \ln p_l) - \frac{(\beta_j + \eta_i' z) \lambda_i}{b(p) c(p, z)} \left[\ln \left(\frac{m}{\overline{m_0}(z) a(p)} \right) \right]^2 \Bigg) \end{aligned} \tag{3.19}$$

where δ_{ij} is the Kronecker delta. The compensated price elasticity can be obtained by using the standard Slutsky equation

$$\epsilon_{ij}^C = \epsilon_{ij} + \mu_i w_j.$$

We estimate both AIDS and QUAIDS models for consumer expenditure on these three consumer goods and services, and also augment these models with demographic characteristics of the sample.

3.5 Results and Discussion

Flexible demand system models like AIDS and QAIDS models contain a large number of parameters. The full parameter estimates, standard errors and p-values obtained from estimating the basic AIDS model, the AIDS model augmented with demographic characteristics, the basic QAIDS model and the QAIDS model augmented with demographic characteristics, using both an unweighted and weighted sample, are at the end of this chapter, on Tables 3.4-3.11. These results are reported primarily for completeness. Recall that the SHS employed a complex survey design that includes uneven respondent selection probabilities over time, which implies that parameter estimation and variance calculations using the weighted sample over years for certain subgroups or variables may require extra caution. Because of this, we report results from both an unweighted and weighed sample. The results are qualitatively identical, so we focus on the unweighted sample results. Note that we do not report the parameter estimates from the year and province indicator variables since their effects are not our primary interest here. These results are available by request from the authors.

In general, the parameters of all models are precisely estimated, as the p-values are quite small. The key parameters of interest in the demand system models are the γ_{ij}'s which indicate the impact of change in price of good j on the budget share of good i. On Tables 3.4-3.11, these parameter estimates are all statistically different from zero at conventional significance levels. Note that the results from the AIDS and QAIDS models, the inclusion of the demographic variables, and the use of weighted and unweighted data has little effect on the sign and significance of the key parameter estimates of interest; the results are robust to these different model specifications.

We focus primarily on the estimated income and price elasticities, since these have the most straightforward economic interpretation. The estimated income and price elasticities are shown on Table 3.2 for the base AIDS model and the AIDS model augmented with demographic characteristics. The results for the QAIDS model, along with the weighted sample results for all models, are found on Tables 3.12, 3.13 and 3.14 at the end of this chapter. The results are quite similar in terms of estimated parameter signs and significance across all models.

The first row on Table 3.2 contains the expenditure elasticity estimate for good or service j. Element in row i, column j in the second panel contains the compensated or Hicksian cross price elasticity of expenditure category i with respect to the changes in the price of good or service j holding utility and prices constant. Element in row i, column j in the third panel contains the uncompensated or Marshallian cross price elasticity of expenditure category i with respect to the changes in the price of good or service j holding income and prices constant. The estimates on the left side of Table 3.2 are for the base AIDS model and the estimates on the right side are for the AIDS model augmented with variables reflecting household characteristics.

On Table 3.2, "Attendance" refers to consumer spending on attendance at live sporting events; "Exercise" refers to consumer spending on leisure time physical activity and exercise; and "Lottery" refers to consumer spending on government sponsored lotteries, which includes sports betting. In each 3×3 cell, the diagonal elements are the estimated own price elasticities and the off-diagonal elements are the estimated cross price elasticities.

The estimated expenditure elasticities are all positive, suggesting that expenditure shares of all three types of consumer goods and services increase with total expenditure. All three are normal goods. The expenditure elasticities for participation in leisure time physical activity are relatively large.

The estimated own price elasticities are uniformly negative, as predicted by standard consumer theory. As the price of each good or service increases, consumer spending on that good or service falls, other things equal. This result holds across all model specifications, including the base and augmented AIDS and QAIDS models for the unweighted and weighted samples. Spending on leisure time participation in physical activity has the highest own price elasticity, and the own price elasticity on attendance at live sporting events is also relatively large. The estimated own price

Table 3.2: Estimated Income and Price Elasticities – AIDS Models

	AIDS Model			Augmented AIDS Model		
	Attendance	Exercise	Lottery	Attendance	Exercise	Lottery
Expenditure Elasticity	1.023	1.287	0.801	0.971	1.253	0.834
Compensated Own and Cross Price Elasticities						
Attendance	-2.902	3.159	-0.257	-3.185	3.263	-0.078
Exercise	0.875	-1.699	0.824	0.906	-2.036	1.130
Lottery	-0.048	0.560	-0.512	-0.016	0.769	-0.753
Uncompensated Own and Cross Price Elasticities						
Attendance	-3.005	2.787	-0.805	-3.283	2.909	-0.598
Exercise	0.745	-2.167	0.135	0.770	-2.492	0.459
Lottery	-0.129	0.269	-0.940	-0.100	0.466	-1.199

elasticity on sports betting is relatively inelastic.

The estimated cross price elasticities of demand contain several interesting features. The uncompensated and compensated estimated cross price elasticity between the price of sports betting and spending on live attendance is negative and relatively small (-0.26); these types of consumer spending are complements, and as the price of sports betting decreases, demand for sports betting and live attendance at sporting events increases, other things equal. This implies that an expansion of legal access to sports betting, which should reduce the price of sports betting, will also increase consumer spending of live attendance at sporting events, which will increase the revenues of sports teams, other things equal. This differs significantly from the claims of sports leagues, who oppose expansions of sports betting because they expect this will reduce their revenues. The opposite estimated cross price elasticities, the compensated and uncompensated cross price elasticity between the price of live attendance at sporting events and consumer spending on sports betting also suggests complementarities exist.

However, the effect of increased expansion of access to sports betting does not have unambiguously positive effects. The uncompensated and compensated estimated cross price elasticity between the price of sports betting and consumer spending on leisure time physical activity is positive; these types of consumer spending are substitutes, and as the price of sports betting decreases, spending on sports betting increases and demand for leisure time participation in paid physical activity decreases, other things equal. In response to a decrease in the price of sports betting, consumers become less physically active and more sedentary, in that they spend less on leisure time participation in physical activity. The opposite estimated cross price elasticities again show the same relationship: they are substitutes.

Humphreys et al. (2014) show that participation in leisure time physical activity in Canada generates better individual health outcomes. Since physically inactive people are less healthy, spending on their health will increase, driving up national health care costs. This is an unintended consequence of expanding legal access to sports betting.

Finally, the uncompensated and compensated estimated cross price elasticity between the price of attending live sporting events and spending on leisure time physical activity is positive; these types of consumer spending are substitutes, and as the price of attending live sporting events decreases, demand for paid leisure time physical activity decreases, other things equal.

This too has important policy implications. Policy makers often claim that hosting sporting events like the Olympic Games will increase participation in physical activity in the local population, providing lasting health and economic benefits from a healthier population. Unfortunately this does not appear to be the case. Bauman et al. (2014) found no evidence that participation in leisure time physical activity in Australia increased following the 2000 Sydney Summer Olympic Games. Smith et al. (2014) report evidence of declines in adolescent participation in physical activity following the 2012 London Summer Olympic Games. In a survey, Weed et al. (2015) find no support for a relationship between hosting the Games and participation in physical activity from a group of 21 different sources. These results provide an economic explanation for why no "demonstration effect" occurs after hosting the Olympic Games, which clearly increases local spending on attending live sporting events.

3.6 Conclusions

This paper develops evidence that consumer spending on sports betting, available through lottery-based games in Canada, and spending on attendance at live sporting events are complements. When the price of one of these goods decreases, consumer spending on both increases. This relationship suggests that increasing access to legal sports betting, which can be interpreted as an increase in the supply of sports betting opportunities, should decrease the price of sports betting, and increase consumer spending on both sports betting and attendance at live sporting events. This evidence contradicts the claims made by professional sports leagues and the NCAA when opposing the expansion of legal sports betting in both Canada and the United States. If these results can be applied to proposed increases in access to legal sports betting in North America, professional and intercollegiate athletic teams and leagues in both countries could expect to see an increase in their revenues following the expansion of legal sports betting opportunities.

Sports teams and leagues have also raised the issue that increased access to legal sports betting will damage the public's perception of the legitimacy of the games sponsored by these leagues by increasing the incentive for gamblers to influence the outcomes of the games by fixing games or matches. Existing evidence suggests that a significant amount of illegal sports betting already takes place in North America, and gamblers have access to on-line book makers. The presence of these alternatives would seem to already provide an incentive for gamblers to fix games or matches. Also Forrest and Simmons (2003) pointed out that the presence of legal sports betting opportunities actually allows law enforcement officials and sports leagues to detect match fixing more easily than would be possible in the absence of legal sports betting opportunities; after all, bookmakers stand to lose money when game or match outcomes are fixed by gamblers. These factors also argue against the claims made by sports teams and leagues when arguing against the expansion of legal sports betting opportunities.

The paper also develops evidence about a negative consequence of expanding legal opportunities to bet on sports as well as expanding opportunities to attend live sporting events. The results from the AIDS and QAIDS models suggest that household spending on leisure time physical activity is a substitute for both spending on both government sponsored lotteries and sports betting, and spending on attendance at live sporting events. As the price of sports betting and attendance at live sporting events declines, demand for participation in physical activity and exercise also declines, as proxied by household spending on fees (including membership and single usage fees) and dues for sports activities, sports and recreation facilities, and health clubs. Reduced participation in physical activity and exercise can have adverse effects of both health outcomes and worker productivity, suggesting that expanding access to sports betting or live sporting events may have negative consequences on the economy.

The results presented here require some caveats. The empirical results come from a sample of Canadian households that reported a positive level of spending on at least one of the three goods and services analyzed here. An expansion of access to legal sports betting could also have an effect on the spending of consumers who did not purchase any of these three goods or services. The behavior of these consumers could differ systematically from the behavior of the households that make up the sample analyzed here. We have used the price of purchasing a lottery ticket as the price of government sponsored lotteries and sports betting, which was either \$1 or \$2 throughout the sample period. The effective price of lottery gambling and sports betting can also depend on the expected return from these activities, which depend on the odds of winning, the size of the prizes and, in the case of sports betting, the over-round on the posted odds on various game outcomes. Although the own price elasticity estimates are negative and significant, the simple price measure for government sponsored lotteries and sports betting used here may not fully capture the prices faced by consumers making decisions to participate in sports betting.

Table 3.3: Demographic Variable Descriptions

Variable	Abbreviation on Table 3.4-3.11
Housing Characteristics (TG)	
Homeowner without mortgage	(*omitted*)
Homeowner with mortgage	$TG2$
Tenants - regular	$TG3$
Tenants - rent-free	$TG4$
Mixed tenure	$TG5$
Major Income Source (MIS)	
No income	(*omitted*)
Paid employment	$MIS2$
Self-employment	$MIS3$
Investment income	$MIS4$
Government transfer payments	$MIS5$
Miscellaneous	$MIS6$
Household Type (HT)	
Single female	(*omitted*)
Single male	$HT2$
Couple only	$HT3$
Couple with children	$HT4$
Single mother	$HT5$
Single father	$HT6$
Other	$HT7$
Household Employment Status (HES)	
Reference person and spouse not working	(*omitted*)
At least one employed part time	$HES2$
Both employed full time	$HES3$

Table 3.4: Estimation Results – Basic AIDS Model

	Coefficient	Std. Err.	p-value
α_1	-0.281	0.023	<0.001
α_2	0.519	0.036	<0.001
α_3	0.762	0.031	<0.001
β_1	0.002	0.0004	<0.001
β_2	0.104	0.001	<0.001
β_3	-0.107	0.001	<0.001
γ_{11}	-0.202	0.025	<0.001
γ_{21}	0.280	0.028	<0.001
γ_{31}	-0.078	0.005	<0.001
γ_{22}	-0.457	0.033	<0.001
γ_{32}	0.176	0.006	<0.001
γ_{33}	-0.098	0.005	<0.001
Observations	145560		

Table 3.5: Estimation Results – AIDS Model Augmented with Demographic Characteristics

	Coefficient	Std. Err.	p-value
α_1	-0.194	0.063	0.002
α_2	1.164	0.088	<0.001
α_3	0.030	0.064	0.642
β_1	0.024	0.006	<0.001
β_2	0.219	0.010	<0.001
β_3	-0.243	0.010	<0.001
γ_{11}	-0.230	0.069	0.001
γ_{21}	0.293	0.079	<0.001
γ_{31}	-0.063	0.013	<0.001
γ_{22}	-0.543	0.093	<0.001
γ_{32}	0.249	0.019	<0.001
γ_{33}	-0.187	0.014	<0.001
$\eta - TG2_1$	-0.004	0.0004	<0.001
$\eta - TG2_2$	-0.010	0.001	<0.001
$\eta - TG2_3$	0.014	0.001	<0.001
$\eta - TG3_1$	-0.004	0.001	<0.001
$\eta - TG3_2$	-0.023	0.001	<0.001
$\eta - TG3_3$	0.027	0.001	<0.001
$\eta - TG4_1$	-0.003	0.002	0.086
$\eta - TG4_2$	-0.018	0.003	<0.001
$\eta - TG4_3$	0.021	0.003	<0.001
$\eta - TG5_1$	-0.003	0.001	0.002
$\eta - TG5_2$	-0.010	0.003	<0.001
$\eta - TG5_3$	0.013	0.003	<0.001
$\eta - MIS2_1$	-0.009	0.006	0.102
$\eta - MIS2_2$	-0.064	0.010	<0.001
$\eta - MIS2_3$	0.073	0.011	<0.001
$\eta - MIS3_1$	-0.008	0.006	0.172
$\eta - MIS3_2$	-0.055	0.010	<0.001
$\eta - MIS3_3$	0.063	0.011	<0.001
$\eta - MIS4_1$	-0.008	0.006	0.167
$\eta - MIS4_2$	-0.039	0.011	<0.001
$\eta - MIS4_3$	0.047	0.011	<0.001
$\eta - MIS5_1$	-0.014	0.006	0.010
$\eta - MIS5_2$	-0.075	0.010	<0.001
$\eta - MIS5_3$	0.090	0.011	<0.001
$\eta - MIS6_1$	-0.010	0.006	0.076
$\eta - MIS6_2$	-0.039	0.010	<0.001

Continued on next page

Table 3.5: Continued

	Coefficient	Std. Err.	p-value
$\eta - MIS6_3$	0.049	0.011	<0.001
$\eta - HT2_1$	0.017	0.001	<0.001
$\eta - HT2_2$	-0.033	0.001	<0.001
$\eta - HT2_3$	0.016	0.001	<0.001
$\eta - HT3_1$	0.006	0.001	<0.001
$\eta - HT3_2$	-0.020	0.002	<0.001
$\eta - HT3_3$	0.015	0.001	<0.001
$\eta - HT4_1$	0.006	0.001	<0.001
$\eta - HT4_2$	-0.014	0.001	<0.001
$\eta - HT4_3$	0.008	0.001	<0.001
$\eta - HT5_1$	0.004	0.001	<0.001
$\eta - HT5_2$	0.007	0.002	<0.001
$\eta - HT5_3$	-0.011	0.002	<0.001
$\eta - HT6_1$	0.015	0.001	<0.001
$\eta - HT6_2$	-0.019	0.003	<0.001
$\eta - HT6_3$	0.005	0.004	0.191
$\eta - HT7_1$	0.007	0.001	<0.001
$\eta - HT7_2$	-0.031	0.001	<0.001
$\eta - HT7_3$	0.024	0.001	<0.001
$\eta - urban_1$	0.002	0.0004	<0.001
$\eta - urban_2$	0.018	0.001	<0.001
$\eta - urban_3$	-0.020	0.001	<0.001
$\eta - age_1$	-0.0004	0.00002	<0.001
$\eta - age_2$	-0.001	0.00004	<0.001
$\eta - age_3$	0.002	0.00004	<0.001
$\eta - HES2_1$	0.0004	0.001	0.582
$\eta - HES2_2$	-0.008	0.001	<0.001
$\eta - HES2_3$	0.007	0.001	<0.001
$\eta - HES3_1$	0.003	0.001	<0.001
$\eta - HES3_2$	-0.0001	0.002	0.972
$\eta - HES3_3$	-0.003	0.002	0.068
$\eta - income_1$	0.00004	2.76e-06	<0.001
$\eta - income_2$	0.0001	6.23e-06	<0.001
$\eta - income_3$	-0.0002	6.23e-06	<0.001
$\rho - TG2$	-0.355	0.079	<0.001
$\rho - TG3$	-0.379	0.080	<0.001
$\rho - TG4$	-0.331	0.096	0.001
$\rho - TG5$	-0.326	0.103	0.002
$\rho - MIS2$	-0.548	0.107	<0.001
$\rho - MIS3$	-0.548	0.117	<0.001

Continued on next page

Table 3.5: Continued

	Coefficient	Std. Err.	p-value
$\rho - MIS4$	-0.580	0.149	<0.001
$\rho - MIS5$	-0.537	0.110	<0.001
$\rho - MIS6$	-0.344	0.148	0.020
$\rho - HT2$	0.001	0.023	0.976
$\rho - HT3$	0.472	0.100	<0.001
$\rho - HT4$	0.174	0.049	<0.001
$\rho - HT5$	-0.011	0.029	0.711
$\rho - HT6$	0.022	0.121	0.855
$\rho - HT7$	-0.021	0.019	0.271
$\rho - urban$	0.102	0.026	<0.001
$\rho - age$	0.008	0.002	<0.001
$\rho - HES2$	-0.248	0.063	<0.001
$\rho - HES3$	0.122	0.066	0.063
$\rho - income$	0.003	0.001	<0.001
Observations	145560		

Table 3.6: Estimation Results – Quadratic AIDS Model

	Coefficient	Std. Err.	p-value
α_1	-0.354	0.024	<0.001
α_2	0.721	0.032	<0.001
α_3	0.633	0.027	<0.001
β_1	0.033	0.003	<0.001
β_2	-0.010	0.003	0.004
β_3	-0.024	0.004	<0.001
γ_{11}	-0.205	0.025	<0.001
γ_{21}	0.284	0.028	<0.001
γ_{31}	-0.079	0.005	<0.001
γ_{22}	-0.424	0.033	<0.001
γ_{32}	0.140	0.007	<0.001
γ_{33}	-0.061	0.005	<0.001
λ_1	-0.003	0.0003	<0.001
λ_2	0.013	0.0003	<0.001
λ_3	-0.009	0.0003	<0.001
Observations	145560		

Table 3.7: Estimation Results – Quadratic AIDS with Demographic Characteristics

	Coefficient	Std. Err.	p-value
α_1	-0.365	0.074	<0.001
α_2	1.513	0.095	<0.001
α_3	-0.148	0.060	0.013
β_1	0.046	0.006	<0.001
β_2	0.182	0.012	<0.001
β_3	-0.228	0.012	<0.001
γ_{11}	-0.197	0.079	0.013
γ_{21}	0.282	0.092	0.002
γ_{31}	-0.085	0.016	<0.001
γ_{22}	-0.590	0.107	<0.001
γ_{32}	0.308	0.020	<0.001
γ_{33}	-0.223	0.013	<0.001
λ_1	-0.004	0.0003	<0.001
λ_2	0.012	0.0003	<0.001
λ_3	-0.007	0.0003	<0.001
$\eta - TG2_1$	-0.0005	0.0004	0.233
$\eta - TG2_2$	-0.021	0.002	<0.001
$\eta - TG2_3$	0.021	0.001	<0.001
$\eta - TG3_1$	0.001	0.0005	0.140
$\eta - TG3_2$	-0.034	0.002	<0.001
$\eta - TG3_3$	0.034	0.002	<0.001
$\eta - TG4_1$	0.001	0.001	0.460
$\eta - TG4_2$	-0.026	0.004	<0.001
$\eta - TG4_3$	0.025	0.004	<0.001
$\eta - TG5_1$	0.0001	0.001	0.863
$\eta - TG5_2$	-0.021	0.003	<0.001
$\eta - TG5_3$	0.021	0.003	<0.001
$\eta - MIS2_1$	-0.008	0.006	0.161
$\eta - MIS2_2$	-0.104	0.012	<0.001
$\eta - MIS2_3$	0.112	0.013	<0.001
$\eta - MIS3_1$	-0.008	0.006	0.163
$\eta - MIS3_2$	-0.092	0.012	<0.001
$\eta - MIS3_3$	0.099	0.013	<0.001
$\eta - MIS4_1$	-0.010	0.006	0.074
$\eta - MIS4_2$	-0.056	0.013	<0.001
$\eta - MIS4_3$	0.066	0.014	<0.001
$\eta - MIS5_1$	-0.013	0.006	0.022

Continued on next page

Table 3.7: Continued

	Coefficient	Std. Err.	p-value
$\eta - MIS5_2$	-0.107	0.012	<0.001
$\eta - MIS5_3$	0.120	0.013	<0.001
$\eta - MIS6_1$	-0.011	0.006	0.049
$\eta - MIS6_2$	-0.072	0.012	<0.001
$\eta - MIS6_3$	0.083	0.013	<0.001
$\eta - HT2_1$	0.016	0.001	<0.001
$\eta - HT2_2$	-0.029	0.001	<0.001
$\eta - HT2_3$	0.013	0.001	<0.001
$\eta - HT3_1$	0.008	0.001	<0.001
$\eta - HT3_2$	-0.026	0.001	<0.001
$\eta - HT3_3$	0.019	0.001	<0.001
$\eta - HT4_1$	0.005	0.001	<0.001
$\eta - HT4_2$	-0.001	0.002	0.449
$\eta - HT4_3$	-0.004	0.001	0.004
$\eta - HT5_1$	0.001	0.001	0.135
$\eta - HT5_2$	0.022	0.003	<0.001
$\eta - HT5_3$	-0.023	0.002	<0.001
$\eta - HT6_1$	0.013	0.001	<0.001
$\eta - HT6_2$	-0.014	0.004	<0.001
$\eta - HT6_3$	0.001	0.003	0.821
$\eta - HT7_1$	0.009	0.001	<0.001
$\eta - HT7_2$	-0.025	0.002	<0.001
$\eta - HT7_3$	0.017	0.001	<0.001
$\eta - urban_1$	-6.41e-06	0.0003	0.984
$\eta - urban_2$	0.015	0.001	<0.001
$\eta - urban_3$	-0.015	0.001	<0.001
$\eta - age_1$	-0.0002	0.00001	<0.001
$\eta - age_2$	-0.001	0.00004	<0.001
$\eta - age_3$	0.001	0.00004	<0.001
$\eta - HES2_1$	0.004	0.001	<0.001
$\eta - HES2_2$	-0.022	0.003	<0.001
$\eta - HES2_3$	0.017	0.002	<0.001
$\eta - HES3_1$	0.006	0.001	<0.001
$\eta - HES3_2$	-0.022	0.003	<0.001
$\eta - HES3_3$	0.016	0.003	<0.001
$\eta - income_1$	0.00001	1.59e-06	<0.001
$\eta - income_2$	0.0001	4.92e-06	<0.001
$\eta - income_3$	-0.0001	5.09e-06	<0.001
$\rho - TG2$	-0.153	0.026	<0.001
$\rho - TG3$	-0.160	0.027	<0.001

Continued on next page

Table 3.7: Continued

	Coefficient	Std. Err.	p-value
$\rho - TG4$	-0.146	0.029	<0.001
$\rho - TG5$	-0.151	0.027	<0.001
$\rho - MIS2$	-0.464	0.087	<0.001
$\rho - MIS3$	-0.453	0.089	<0.001
$\rho - MIS4$	-0.186	0.171	0.277
$\rho - MIS5$	-0.368	0.099	<0.001
$\rho - MIS6$	-0.413	0.096	<0.001
$\rho - HT2$	-0.002	0.002	0.348
$\rho - HT3$	0.009	0.002	<0.001
$\rho - HT4$	0.077	0.015	<0.001
$\rho - HT5$	0.055	0.017	0.001
$\rho - HT6$	-0.006	0.015	0.685
$\rho - HT7$	0.011	0.003	<0.001
$\rho - urban$	-0.004	0.004	0.355
$\rho - age$	0.003	0.001	<0.001
$\rho - HES2$	-0.413	0.078	<0.001
$\rho - HES3$	-0.412	0.077	<0.001
$\rho - income$	0.0001	0.00003	0.053
Observations	145560		

Table 3.8: Estimation Results – Basic AIDS Model (Weighted Sample)

	Coefficient	Robust Std. Err.	p-value
α_1	-0.366	0.036	$<$0.001
α_2	0.402	0.054	$<$0.001
α_3	0.964	0.044	$<$0.001
β_1	0.008	0.001	$<$0.001
β_2	0.112	0.001	$<$0.001
β_3	-0.120	0.001	$<$0.001
γ_{11}	-0.295	0.036	$<$0.001
γ_{21}	0.379	0.042	$<$0.001
γ_{31}	-0.084	0.008	$<$0.001
γ_{22}	-0.566	0.049	$<$0.001
γ_{32}	0.187	0.009	$<$0.001
γ_{33}	-0.103	0.006	$<$0.001
Observations	145560		
Weighted Sample Size	1.22e+08		

Table 3.9: Estimation Results – Augments AIDS Model (Weighted Sample)

	Coefficient	Robust Std. Err.	p-value
α_1	-0.280	0.102	0.006
α_2	1.420	0.129	<0.001
α_3	-0.140	0.091	0.123
β_1	0.037	0.011	0.001
β_2	0.218	0.018	<0.001
β_3	-0.254	0.015	<0.001
γ_{11}	-0.418	0.106	<0.001
γ_{21}	0.492	0.123	<0.001
γ_{31}	-0.074	0.021	0.001
γ_{22}	-0.817	0.144	<0.001
γ_{32}	0.324	0.028	<0.001
γ_{33}	-0.251	0.020	<0.001
$\eta - TG2_1$	-0.004	0.001	<0.001
$\eta - TG2_2$	-0.012	0.002	<0.001
$\eta - TG2_3$	0.016	0.002	<0.001
$\eta - TG3_1$	-0.004	0.001	<0.001
$\eta - TG3_2$	-0.027	0.002	<0.001
$\eta - TG3_3$	0.031	0.002	<0.001
$\eta - TG4_1$	-0.006	0.002	0.002
$\eta - TG4_2$	-0.024	0.004	<0.001
$\eta - TG4_3$	0.030	0.004	<0.001
$\eta - TG5_1$	-0.003	0.001	0.058
$\eta - TG5_2$	-0.013	0.004	0.002
$\eta - TG5_3$	0.015	0.004	<0.001
$\eta - MIS2_1$	-0.016	0.011	0.146
$\eta - MIS2_2$	-0.056	0.019	0.003
$\eta - MIS2_3$	0.072	0.016	<0.001
$\eta - MIS3_1$	-0.017	0.011	0.138
$\eta - MIS3_2$	-0.046	0.019	0.016
$\eta - MIS3_3$	0.063	0.016	<0.001
$\eta - MIS4_1$	-0.017	0.011	0.120
$\eta - MIS4_2$	-0.035	0.019	0.069
$\eta - MIS4_3$	0.053	0.017	0.001
$\eta - MIS5_1$	-0.021	0.011	0.059
$\eta - MIS5_2$	-0.062	0.019	0.001
$\eta - MIS5_3$	0.083	0.016	<0.001
$\eta - MIS6_1$	-0.017	0.011	0.137
$\eta - MIS6_2$	-0.034	0.019	0.072

Continued on next page

Table 3.9: Continued

	Coefficient	Robust Std. Err.	p-value
$\eta - MIS6_3$	0.051	0.016	0.002
$\eta - HT2_1$	0.015	0.001	<0.001
$\eta - HT2_2$	-0.030	0.002	<0.001
$\eta - HT2_3$	0.014	0.002	<0.001
$\eta - HT3_1$	0.004	0.001	<0.001
$\eta - HT3_2$	-0.022	0.002	<0.001
$\eta - HT3_3$	0.018	0.002	<0.001
$\eta - HT4_1$	0.005	0.001	<0.001
$\eta - HT4_2$	-0.012	0.002	<0.001
$\eta - HT4_3$	0.006	0.002	0.004
$\eta - HT5_1$	0.004	0.001	0.001
$\eta - HT5_2$	0.010	0.003	0.002
$\eta - HT5_3$	-0.014	0.003	<0.001
$\eta - HT6_1$	0.013	0.002	<0.001
$\eta - HT6_2$	-0.015	0.005	0.005
$\eta - HT6_3$	0.002	0.006	0.683
$\eta - HT7_1$	0.006	0.001	<0.001
$\eta - HT7_2$	-0.032	0.002	<0.001
$\eta - HT7_3$	0.026	0.002	<0.001
$\eta - urban_1$	0.001	0.001	0.047
$\eta - urban_2$	0.012	0.001	<0.001
$\eta - urban_3$	-0.014	0.001	<0.001
$\eta - age_1$	-0.0004	0.00002	<0.001
$\eta - age_2$	-0.001	0.0001	<0.001
$\eta - age_3$	0.002	0.0001	<0.001
$\eta - HES2_1$	-0.001	0.001	0.494
$\eta - HES2_2$	-0.006	0.002	0.003
$\eta - HES2_3$	0.007	0.002	0.001
$\eta - HES3_1$	0.002	0.001	0.060
$\eta - HES3_2$	-0.003	0.002	0.232
$\eta - HES3_3$	0.001	0.002	0.654
$\eta - income_1$	0.00003	4.43e-06	<0.001
$\eta - income_2$	0.0001	0.00001	<0.001
$\eta - income_3$	-0.0001	0.00002	<0.001
$\rho - TG2$	-0.214	0.072	0.003
$\rho - TG3$	-0.317	0.088	<0.001
$\rho - TG4$	-0.313	0.093	0.001
$\rho - TG5$	-0.205	0.094	0.029
$\rho - MIS2$	-0.637	0.111	<0.001
$\rho - MIS3$	-0.624	0.118	<0.001

Continued on next page

Table 3.9: Continued

	Coefficient	Robust Std. Err.	p-value
$\rho - MIS4$	-0.655	0.110	<0.001
$\rho - MIS5$	-0.647	0.110	<0.001
$\rho - MIS6$	-0.616	0.117	<0.001
$\rho - HT2$	-0.035	0.019	0.061
$\rho - HT3$	0.087	0.038	0.021
$\rho - HT4$	0.211	0.063	0.001
$\rho - HT5$	0.156	0.064	0.015
$\rho - HT6$	0.027	0.112	0.812
$\rho - HT7$	-0.035	0.019	0.063
$\rho - urban$	0.018	0.011	0.097
$\rho - age$	0.007	0.002	0.001
$\rho - HES2$	-0.156	0.057	0.006
$\rho - HES3$	-0.021	0.049	0.662
$\rho - income$	0.001	0.0004	0.086
Observations	145560		
Weighted Sample Size	1.22e+08		

Table 3.10: Estimation Results – Quadratic AIDS Model (Weighted Sample)

	Coefficient	Robust Std. Err.	p-value
α_1	-0.470	0.038	<0.001
α_2	0.638	0.048	<0.001
α_3	0.832	0.041	<0.001
β_1	0.049	0.004	<0.001
β_2	-0.017	0.005	0.001
β_3	-0.032	0.005	<0.001
γ_{11}	-0.300	0.036	<0.001
γ_{21}	0.385	0.042	<0.001
γ_{31}	-0.084	0.008	<0.001
γ_{22}	-0.516	0.050	<0.001
γ_{32}	0.132	0.010	<0.001
γ_{33}	-0.047	0.007	<0.001
λ_1	-0.004	0.0004	<0.001
λ_2	0.013	0.0004	<0.001
λ_3	-0.009	0.0004	<0.001
Observations	145560		
Weighted Sample Size	1.22e+08		

Table 3.11: Estimation Results – Augmented Quadratic AIDS Model (Weighted Sample)

	Coefficient	Robust Std. Err.	p-value
α_1	-0.152	0.105	0.146
α_2	1.401	0.140	<0.001
α_3	-0.249	0.105	0.018
β_1	0.057	0.013	<0.001
β_2	0.183	0.020	<0.001
β_3	-0.240	0.018	<0.001
γ_{11}	-0.267	0.108	0.013
γ_{21}	0.298	0.126	0.018
γ_{31}	-0.031	0.022	0.163
γ_{22}	-0.591	0.146	<0.001
γ_{32}	0.294	0.030	<0.001
γ_{33}	-0.263	0.023	<0.001
λ_1	-0.004	0.0004	<0.001
λ_2	0.012	0.0004	<0.001
λ_3	-0.008	0.0004	<0.001
$\eta - TG2_1$	-0.001	0.001	0.005
$\eta - TG2_2$	-0.020	0.003	<0.001
$\eta - TG2_3$	0.022	0.002	<0.001
$\eta - TG3_1$	0.00003	0.001	0.959
$\eta - TG3_2$	-0.037	0.003	<0.001
$\eta - TG3_3$	0.037	0.003	<0.001
$\eta - TG4_1$	-0.001	0.002	0.468
$\eta - TG4_2$	-0.035	0.005	<0.001
$\eta - TG4_3$	0.036	0.005	<0.001
$\eta - TG5_1$	-0.001	0.001	0.601
$\eta - TG5_2$	-0.023	0.005	<0.001
$\eta - TG5_3$	0.023	0.004	<0.001
$\eta - MIS2_1$	-0.016	0.012	0.196
$\eta - MIS2_2$	-0.091	0.022	<0.001
$\eta - MIS2_3$	0.107	0.020	<0.001
$\eta - MIS3_1$	-0.017	0.012	0.167
$\eta - MIS3_2$	-0.080	0.022	<0.001
$\eta - MIS3_3$	0.098	0.021	<0.001
$\eta - MIS4_1$	-0.018	0.012	0.139
$\eta - MIS4_2$	-0.066	0.022	0.002
$\eta - MIS4_3$	0.085	0.021	<0.001
$\eta - MIS5_1$	-0.020	0.012	0.103

Continued on next page

Table 3.11: Continued

	Coefficient	Robust Std. Err.	p-value
$\eta - MIS5_2$	-0.093	0.021	<0.001
$\eta - MIS5_3$	0.113	0.020	<0.001
$\eta - MIS6_1$	-0.018	0.012	0.147
$\eta - MIS6_2$	-0.066	0.022	0.002
$\eta - MIS6_3$	0.084	0.021	<0.001
$\eta - HT2_1$	0.015	0.001	<0.001
$\eta - HT2_2$	-0.027	0.002	<0.001
$\eta - HT2_3$	0.012	0.002	<0.001
$\eta - HT3_1$	0.006	0.001	<0.001
$\eta - HT3_2$	-0.027	0.002	<0.001
$\eta - HT3_3$	0.021	0.002	<0.001
$\eta - HT4_1$	0.005	0.001	<0.001
$\eta - HT4_2$	-0.007	0.002	0.007
$\eta - HT4_3$	0.002	0.002	0.417
$\eta - HT5_1$	0.001	0.001	0.128
$\eta - HT5_2$	0.018	0.004	<0.001
$\eta - HT5_3$	-0.020	0.004	<0.001
$\eta - HT6_1$	0.011	0.002	<0.001
$\eta - HT6_2$	-0.011	0.006	0.056
$\eta - HT6_3$	0.0001	0.005	0.979
$\eta - HT7_1$	0.009	0.001	<0.001
$\eta - HT7_2$	-0.033	0.002	<0.001
$\eta - HT7_3$	0.024	0.002	<0.001
$\eta - urban_1$	5.42e-06	0.0005	0.991
$\eta - urban_2$	0.011	0.001	<0.001
$\eta - urban_3$	-0.011	0.001	<0.001
$\eta - age_1$	-0.0002	0.00002	<0.001
$\eta - age_2$	-0.001	0.00006	<0.001
$\eta - age_3$	0.001	0.00006	<0.001
$\eta - HES2_1$	0.002	0.001	0.003
$\eta - HES2_2$	-0.025	0.004	<0.001
$\eta - HES2_3$	0.022	0.003	<0.001
$\eta - HES3_1$	0.004	0.001	<0.001
$\eta - HES3_2$	-0.028	0.004	<0.001
$\eta - HES3_3$	0.024	0.004	<0.001
$\eta - income_1$	0.00002	3.15e-06	<0.001
$\eta - income_2$	0.0001	0.00001	<0.001
$\eta - income_3$	-0.0001	0.00001	<0.001
$\rho - TG2$	-0.100	0.036	0.006
$\rho - TG3$	-0.118	0.041	0.004

Continued on next page

Table 3.11: Continued

	Coefficient	Robust Std. Err.	p-value
$\rho - TG4$	-0.119	0.042	0.004
$\rho - TG5$	-0.102	0.039	0.008
$\rho - MIS2$	-0.614	0.125	<0.001
$\rho - MIS3$	-0.609	0.127	<0.001
$\rho - MIS4$	-0.616	0.125	<0.001
$\rho - MIS5$	-0.577	0.135	<0.001
$\rho - MIS6$	-0.617	0.125	<0.001
$\rho - HT2$	-0.001	0.002	0.536
$\rho - HT3$	0.004	0.004	0.232
$\rho - HT4$	0.061	0.021	0.004
$\rho - HT5$	0.050	0.025	0.043
$\rho - HT6$	0.003	0.019	0.893
$\rho - HT7$	-0.001	0.002	0.620
$\rho - urban$	-0.003	0.004	0.525
$\rho - age$	0.002	0.001	0.002
$\rho - HES2$	-0.287	0.097	0.003
$\rho - HES3$	-0.287	0.097	0.003
$\rho - income$	0.00005	0.00006	0.417
Observations	145560		
Weighted Sample Size	1.22e+08		

Table 3.12: Estimated Price and Income Elasticities – QAIDS

	Basic QAIDS			Augmented QAIDS Model		
	Attendance	Exercise	Lottery	Attendance	Exercise	Lottery
Expenditure Elasticity	1.0238	1.2862	0.8011	0.8638	1.3376	0.7962
Compensated Own and Cross Price Elasticities						
Attendance	-2.9165	3.1807	-0.3047	-2.8509	3.1070	-0.4393
Exercise	0.8863	-1.7074	0.8623	0.8910	-2.1513	1.3964
Lottery	-0.0533	0.5615	-0.5285	-0.0688	0.8770	-0.8662
Uncompensated Own and Cross Price Elasticities						
Attendance	-3.0197	2.8082	-0.8529	-2.9379	2.7927	-0.9018
Exercise	0.7567	-2.1753	0.1736	0.7562	-2.6379	0.6803
Lottery	-0.1340	0.2701	-0.9575	-0.1491	0.5873	-1.2925

Table 3.13: Estimated Price and Income Elasticities – AIDS (Weighted Sample)

	Basic AIDS			Augmented AIDS		
	Attendance	Exercise	Lottery	Attendance	Exercise	Lottery
Expenditure Elasticity	1.0809	1.3088	0.7750	1.0409	1.2732	0.8067
Compensated Own and Cross Price Elasticities						
Attendance	-3.8185	4.1930	-0.3745	-5.0510	5.2662	-0.2152
Exercise	1.1638	-1.9340	0.7702	1.4614	-2.7554	1.2940
Lottery	-0.0721	0.5249	-0.4529	-0.0424	0.8810	-0.8387
Uncompensated Own and Cross Price Elasticities						
Attendance	-3.9274	3.7997	-0.9532	-5.1559	4.8875	-0.7725
Exercise	1.0319	-2.4101	0.0694	1.3331	-3.2186	0.6122
Lottery	-0.1502	0.2430	-0.8678	-0.1236	0.5876	-1.2706

Table 3.14: Estimated Price and Income Elasticities – QAIDS Model (Weighted Sample)

	Basic QAIDS			QAIDS w/ Demographics		
	Attendance	Exercise	Lottery	Attendance	Exercise	Lottery
Expenditure Elasticity	1.0960	1.2958	0.7809	0.9158	1.3753	0.7608
Compensated Own and Cross Price Elasticities						
Attendance	-3.8153	4.1992	-0.4192	-3.5391	3.2851	0.0656
Exercise	1.1677	-1.9392	0.8025	0.9423	-2.1337	1.3449
Lottery	-0.0753	0.5273	-0.4664	0.0258	0.8315	-0.9261
Uncompensated Own and Cross Price Elasticities						
Attendance	-3.9258	3.8005	-1.0060	-3.6314	2.9520	-0.4247
Exercise	1.0371	-2.4107	0.1087	0.8037	-2.6341	0.6085
Lottery	-0.1540	0.2432	-0.8845	-0.0509	0.5547	-1.3335

Chapter 4

Concluding Remarks

The three essays of my dissertation apply microeconomic theory to investigate individual's behavior in workplace and consumption market. Utilizing different micro-level datasets, we empirically examine how high-skilled team members make their own efforts in response to teammates' performance; how households' consumption commitments affect their risk preferences; and how consumers allocate their spending to four alternative leisure time activities.

The first essay studies professional sport players' shirking behavior in a team environment, using data from the National Football League. Our results are consistent with the presence of a low powered punishment mechanism through which a group of team members can punish another team member who shirks during their repeated workplace interaction. The results extend the growing literature of the importance of peer effects. This study has important implications for research on the pay and performance relationship. It provides evidence for optimal incentives for a high-skill and repeatedly interacting team.

The second essay extends a recent research on the seemingly contradictory behavior of buying insurance and gambling at the same time. It also makes an empirical contribution to the literature using Canadian household survey data. The paper explores how consumer's characteristics, particularly consumption commitments, affect participation decisions in two kinds of gambling activities and life insurance purchase. The consumption commitments, including housing expenditure, magnify or mitigate households' risk aversion or loving. The results give helpful implications for the development and regulation of gambling and insurance products.

The third essay provides useful inputs to the discussions on the legalization of sports betting in the US and Canada. Specifically, sports leagues often oppose the legalization and claim that it would reduce their revenue. But our analysis suggests that increasing spending on sports betting will increase consumers' spending on spectator sports and in turn increase sports leagues' revenue. On the flip side, increasing spending on sports betting or spectator sports events reduces participation in physical exercise, and may carry with it adverse health impacts. The relationships between those consumer spending have important policy implications for both gambling and health policy.

Bibliography

Abdel-Ghany, M. and Sharpe, D. L. 2001. Lottery expenditures in Canada: Regional analysis of probability of purchase, amount of purchase, and incidence. *Family and Consumer Sciences Research Journal*, 30(1):64–78.

Baltagi, B., Bresson, G., and Pirotte, A. 2003. Fixed effects, random effects or Hausman-Taylor?: A pretest estimator. *Economics Letters*, 79(3):361–369.

Banks, J., Blundell, R., and Lewbel, A. 1997. Quadratic Engel curves and consumer demand. *Review of Economics and Statistics*, 79(4):527–539.

Bauman, A., Bellew, B., and Craig, C. L. 2014. Did the 2000 Sydney Olympics increase physical activity among adult Australians? *British journal of sports medicine*, pages bjsports–2013.

Berri, D. J. and Simmons, R. 2009. Race and the evaluation of signal callers in the National Football League. *Journal of Sports Economics*, 10(1):23–43.

Berri, D. J., Simmons, R., Van Gilder, J., and O'Neill, L. 2011. What does it mean to find the face of the franchise? physical attractiveness and the evaluation of athletic performance. *Economics Letters*, 111(3):200–202.

Blow, L., Lechene, V., and Levell, P. 2012. Using the CE to model household demand. In *Improving the Measurement of Consumer Expenditures*, NBER Chapters. National Bureau of Economic Research.

Buraimo, B., Forrest, D., and Simmons, R. 2008. Insights for clubs from modelling match attendance in football. *Journal of the Operational Research Society*, 60(2):147–155.

Calvet, L. E. and Sodini, P. 2014. Twin picks: Disentangling the determinants of risk-taking in household portfolios. *The Journal of Finance*, 69(2):867–906.

Cappellari, L. and Jenkins, S. P. 2004. Modelling low income transitions. *Journal of Applied Econometrics*, 19(5):593–610.

Chan, T. Y., Li, J., and Pierce, L. 2014. Compensation and peer effects in competing sales teams. *Management Science*, 60(8):1965–1984.

Chang, D. and Serletis, A. 2013. The demand for gasoline: Evidence from household survey data. *Journal of Applied Econometrics*. In press.

Che, Y. and Yoo, S. 2001. Optimal incentives for teams. *American Economic Review*, 91(3):525–541.

Chetty, R. and Szeidl, A. 2007. Consumption commitments and risk preferences. *The Quarterly Journal of Economics*, 122(2):831–877.

Clotfelter, C. and Cook, P. 1991. *Selling Hope: State Lotteries in America*. National Bureau of Economic Research Bill. Harvard University Press.

Conlisk, J. 1993. The utility of gambling. *Journal of Risk and Uncertainty*, 6:255–275.

Deaton, A. and Muellbauer, J. 1980. An almost ideal demand system. *The American economic review*, pages 312–326.

Depken, C. and Haglund, L. 2011. Peer effects in team sports: Empirical evidence from NCAA relay teams. *Journal of Sports Economics*, 12(1):3.

Depken, C. A. 2000. Wage disparity and team productivity: Evidence from Major League Baseball. *Economics Letters*, 67(1):87–92.

Diecidue, E., Schmidt, U., and Wakker, P. 2004. The utility of gambling reconsidered. *Journal of Risk and Uncertainty*, 29:241–259.

Dixit, K. and Pal, R. 2010. The impact of group incentives on performance of small firms: Hausman–Taylor estimates. *Managerial and Decision Economics*, 31(6):403–414.

Dowell, R. 1985. Risk preference and the work-leisure trade-off. *Economic Inquiry*, 23(4):691–701.

Dowell, R. S. and McLaren, K. R. 1986. An intertemporal analysis of the interdependence between risk preference, retirement, and work rate decisions. *Journal of Political Economy*, 94(3):667–82.

Eden, B. 1977. The role of insurance and gambling in allocating risk over time. *Journal of Economic Theory*, 16(2):228–246.

Eden, B. 1980. The insurance-buying gambler. *Economic Inquiry*, 18(3):504–508.

Farrell, L. and Walker, I. 1999. The welfare effects of lotto: Evidence from the UK. *Journal of Public Economics*, 72(1):99–120.

Forrest, D. and Pérez, L. 2011. Football pools and lotteries: Substitute roads to riches? *Applied Economics Letters*, 18(13):1253–1257.

Forrest, D. and Simmons, R. 2003. Sport and gambling. *Oxford Review of Economic Policy*, 19(4):598–611.

Fort, R. and Quirk, J. 1995. Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature*, 33(3):1265–1299.

Friedman, M. and Savage, L. 1948. The utility analysis of choices involving risk. *The Journal of Political Economy*, 56(4):279–304.

García, J., Pérez, L., and Rodríguez, P. 2008. Football pool sales: How important is a football club in the top division? *International Journal of Sport Finance*, 3(3):167–176.

García, J. and Rodríguez, P. 2007. The demand for football pools in Spain The role of price, prizes, and the composition of the coupon. *Journal of Sports Economics*, 8(4):335–354.

Gebreegziabher, Z., Mekonnen, A., Kassie, M., and Köhlin, G. 2012. Urban energy transition and technology adoption: The case of Tigrai, northern Ethiopia. *Energy Economics*, 34(2):410–418.

Geweke, J. F., Keane, M. P., and Runkle, D. E. 1997. Statistical inference in the multinomial multiperiod probit model. *Journal of Econometrics*, 80(1):125–165.

Greene, W. 2004. Convenient estimators for the panel probit model: Further results. *Empirical Economics*, 29(1):21–47.

Greene, W. H. 2008. *Econometric analysis*, volume 6th ed. Prentice hall Upper Saddle River, NJ.

Guryan, J., Kroft, K., and Notowidigdo, M. 2009. Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics*, 1(4):34–68.

Guta, D. D. 2012. Application of an almost ideal demand system (AIDS) to Ethiopian rural residential energy use: Panel data evidence. *Energy Policy*, 50:528–539.

Hajivassiliou, V., McFadden, D., and Ruud, P. 1996. Simulation of multivariate normal rectangle probabilities and their derivatives theoretical and computational results. *Journal of econometrics*, 72(1):85–134.

Hausman, J. A. and Taylor, W. E. 1981. Panel data and unobservable individual effects. *Econometrica*, 49(6):1377–1398.

Hendricks, W., DeBrock, L., and Koenker, R. 2003. Uncertainty, hiring, and subsequent performance: The NFL draft. *Journal of Labor Economics*, 21(4).

Holmstrom, B. 1982. Moral hazard in teams. *The Bell Journal of Economics*, 13(2):324–340.

Humphreys, B. R., McLeod, L., and Ruseski, J. E. 2014. Physical activity and health outcomes: Evidence from Canada. *Health economics*, 23(1):33–54.

Humphreys, B. R. and Perez, L. 2012. Network externalities in consumer spending on lottery games: Evidence from Spain. *Empirical Economics*, 42:929–945.

Hyslop, D. R. 1999. State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. *Econometrica*, 67(6):1255–1294.

Ichino, A. and Maggi, G. 2000. Work environment and individual background: Explaining regional shirking differentials in a large Italian firm. *The Quarterly Journal of Economics*, 115(3):1057–1090.

Ishida, J. 2006. Team incentives under relative performance evaluation. *Journal of Economics & Management Strategy*, 15(1):187–206.

Jones, L. E. 2008. A note on the joint occurrence of insurance and gambling. *Macroeconomic Dynamics*, 12:97–111.

Kahn, L. 1993. Free agency, long-term contracts and compensation in Major League Baseball: Estimates from panel data. *The Review of Economics and Statistics*, 75(1):157–164.

Kahneman, D. and Tversky, A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291.

Kandel, E. and Lazear, E. P. 1992. Peer pressure and partnerships. *Journal of Political Economy*, 100(4):801–17.

Keefer, Q. 2011. Compensation discrimination for defensive players: Applying quantile regression to the National Football League market for linebackers. *Journal of Sports Economics*.

Kim, Y. 1973. Choice in the lottery-insurance situation augmented-income approach. *The Quarterly Journal of Economics*, 87(1):148–156.

Koohi-Kamali, F. 2013. Estimation of equivalence scales under convertible rationing. *Review of Income and Wealth*, 59(1):113–132.

Kvaløy, O. and Olsen, T. E. 2012. The rise of individual performance pay. *Journal of Economics & Management Strategy*, 21(2):493–518.

Kwang, N. Y. 1965. Why do people buy lottery tickets? Choices involving risk and the indivisibility of expenditure. *The Journal of Political Economy*, pages 530–535.

Lazear, E. P. 1989. Pay equality and industrial politics. *Journal of Political Economy*, 97(3):561–80.

Mangion, M.-L., Cooper, C., Cortes-Jimenez, I., and Durbarry, R. 2012. Measuring the effect of subsidization on tourism demand and destination competitiveness through the AIDS model: An evidence-based approach to tourism policymaking. *Tourism Economics*, 18(6):1251–1272.

Manski, C. F. 1993. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542.

Markowitz, H. 1952. The utility of wealth. *The Journal of Political Economy*, pages 151–158.

Mas, A. and Moretti, E. 2009. Peers at work. *American Economic Review*, 99(1):112–145.

Mundlak, Y. 1978. On the pooling of time series and cross section data. *Econometrica*, 46(1):69–85.

Neckermann, S., Cueni, R., and Frey, B. S. 2014. Awards at work. *Labour Economics*. Forthcoming.

Nyman, J., Welte, J., and Dowd, B. 2008. Something for nothing: A model of gambling behavior. *Journal of Socio-Economics*, 37(6):2492–2504.

Persico, N., Postlewaite, A., and Silverman, D. 2004. The effect of adolescent experience on labor market outcomes: The case of height. *The Journal of Political Economy*, 112(5):1019–1053.

Poi, B. P. 2012. Easy demand-system estimation with quaids. *Stata Journal*, 12(3):433–446(14).

Ray, R. 1983. Measuring the costs of children: An alternative approach. *Journal of Public Economics*, 22(1):89–102.

Sauer, R. D. 2001. The political economy of gambling regulation. *Managerial and Decision Economics*, 22(1-3):5–15.

Schultz, T. 2002. Wage gains associated with height as a form of health human capital. *The American Economic Review*, 92(2):349–353.

Scott, F. and Garen, J. 1994. Probability of purchase, amount of purchase, and the demographic incidence of the lottery tax. *Journal of Public Economics*, 54(1):121 – 143.

Simmons, R. 2007. Prohibition of gambling. In Meadowcroft, J., editor, *Prohibitions (Readings in Economics)*. Institute of Economic Affairs.

Simmons, R. and Berri, D. 2009. Gains from specialization and free agency: The story from the gridiron. *Review of Industrial Organization*, 34(1):81–98.

Smith, N., Lewis, D., Fahy, A., Thompson, C., Clark, C., Stansfeld, S., Cummins, S., Taylor, S., Eldridge, S., Greenhalgh, T., et al. 2014. Changes in physical activity in East London's adolescents following the 2012 Olympic Games: Findings from the prospective Olympic Regeneration in East London (ORiEL) cohort study. *Journal of Epidemiology and Community Health*, 68(Suppl 1):A23–A24.

Strumpf, K. 2003. Illegal sports bookmakers. University of North Carolina Chapel Hill Working Paper.

Szpiro, G. 1992. Insurance buying gamblers. *Theory and decision*, 32(2):203–207.

Tversky, A. and Kahneman, D. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

Wakker, P. P. 2010. *Prospect theory: For risk and ambiguity*. Cambridge University Press.

Weed, M., Coren, E., Fiore, J., Wellard, I., Chatziefstathiou, D., Mansfield, L., and Dowse, S. 2015. The Olympic Games and raising sport participation: A systematic review of evidence and an interrogation of policy for a demonstration effect. *European Sport Management Quarterly*, 15(2):195–226.

Wilde, J. 2000. Identification of multiple equation probit models with endogenous dummy regressors. *Economics letters*, 69(3):309–312.

Worthington, A., Brown, K., Crawford, M., and Pickernell, D. 2007. Gambling participation in Australia: Findings from the national Household Expenditure Survey. *Review of Economics of the Household*, 5:209–221.