

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Essays in behavioral economics : evidence from the field

Permalink

<https://escholarship.org/uc/item/2pz9q34k>

Author

Lien, Jaimie Wei-hung

Publication Date

2010

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays in Behavioral Economics: Evidence from the Field

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor
of Philosophy

in

Economics

by

Jaimie Wei-hung Lien

Committee in charge:

Professor Vincent P. Crawford, Co-Chair

Professor Julie Berry Cullen, Co-Chair

Professor Uri Gneezy

Professor Craig McIntosh

Professor Uri Simonsohn

Professor Joel Sobel

2010

Copyright

Jaimie Wei-hung Lien, 2010

All rights reserved.

The Dissertation of Jaimie Wei-hung Lien is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Co-Chair

University of California, San Diego

2010

DEDICATION

Dedicated to my parents and my brother,
my grandparents, and my extended family.

EPIGRAPH

I give myself very good advice
But I very seldom follow it
That explains the trouble that I'm always in
Be patient
Is very good advice
But the waiting makes me curious
And I'd love the change
Should something strange begin

Well, I went along my merry way
And I never stopped to reason
I should have known there'd be a price to pay
Someday, someday

I give myself very good advice
But I very seldom follow it
Will I ever learn to do the things I should?

- *"Alice in Wonderland"* (1951)

TABLE OF CONTENTS

Signature Page.....	iii
Dedication.....	iv
Epigraph.....	v
Table of Contents.....	vi
List of Figures	x
List of Tables	xii
Acknowledgements.....	xiv
Vita.....	xvi
Abstract of the Dissertation	xvii
Chapter 1: Spending and Learning in a University Dining Program.....	1
1.1 Introduction	1
1.2 The Meal Plan	5
1.3 Discussion of smoothing motives	10
1.3.1 Censoring and selection effects	11
1.3.2 Budgetary and taste heterogeneity.....	13
1.3.3 Meal point trading	15
1.4 Proposed models	15
1.4.1 School's recommended plan.....	16
1.4.2 Complete smoothing.....	17
1.4.3 Biased smoothing	19
1.4.4 Biased smoothing with learning	21
1.5 Data and preliminary analysis	22

1.5.1 Censoring the sample.....	22
1.5.2 Coordination of start-date	23
1.5.3 Preliminary data analysis	24
1.6 Estimation.....	29
1.6.1 Sorting of students by estimation results	31
1.7 Results	33
1.7.1 Complete smoothing.....	33
1.7.2 Biased smoothing	36
1.7.3 Biased smoothing with learning	39
1.7.4 School's suggestions.....	43
1.8 Conclusions	44
1.8.1 Policy implications for meal plans	47
1.8.2 General policy insights	50
Appendix 1.A: Survey responses	52
References:	65
Chapter 2: Slot Machine Stopping Decisions: Evidence for Prospect Theory Preferences?	66
2.1 Introduction	66
2.2 Literature review	70
2.3 Models of decisions under uncertainty	73
2.3.1 Expected utility.....	73
2.3.2 Prospect theory	75
2.3.3 Predictions of PT and EU for slot machine gambling	78

2.4 Data	80
2.5 Predictions for individual stopping decisions	84
2.5.1 Normality testing	89
2.6 Results	94
2.7 Conclusions	100
Appendix 2.A:	104
Appendix 2.B:	105
Appendix 2.C:	109
References:	112
Chapter 3: Impatiently Saving for Leisure: Layaway as a Loose Commitment Savings Contract.....	114
3.1 Introduction	114
3.2 Layaway contracts.....	116
3.2.1 Historical context.....	116
3.2.2 Online layaway rules	117
3.3 Theoretical motivation	121
3.4 Data	123
3.5 Results	127
3.5.1 Contract choice	127
3.5.2 Default and cancellation	132
3.6 Conclusions	133
Appendix 3.A: Categorization of items.....	136
Appendix 3.B: Commission effects	138

References:	141
-------------------	-----

LIST OF FIGURES

Figure 1.1.1: Aggregate expenditure volume	24
Figure 1.1.2: Remaining balances at end-of-year	25
Figure 1.1.3: Remaining balances (greater than 100 only) at end-of-year	25
Figure 1.2: Meal distributions, by participant	27
Figure 1.3: Retail purchases, by week	29
Figure 1.4: Examples of complete smoothing individuals	35
Figure 1.5: Examples of biased smoothing individuals	37
Figure 1.6: Biased smoothing types, distribution of fractions λ	38
Figure 1.7: Examples of biased smoothing with learning individuals	40
Figure 1.8: Biased smoothing with learning types, fraction of optimal plan	42
Figure 1.9: Biased smoothing with learning types, week of learning	43
Figure 2.1: Example neoclassical utility function	74
Figure 2.2: Example Prospect Theory value function	76
Figure 2.3: Example Prospect Theory value function with diminishing sensitivity.....	78
Figure 2.4a:	82
Figure 2.4b:	82
Figure 2.5: Distribution of winning levels at quitting	84
Figure 2.6: Varying kurtosis	87
Figure 2.7: Skewness	88
Figure 2.8a:	90
Figure 2.8b:	90
Figure 2.9: Histograms by bins	97
Figure 2.10: Q-Q plots	99

Figure 3.1: Replica of online layaway calculator.....	119
Figure 3.2: Distribution of number of purchases/contracts, by customer.....	127

LIST OF TABLES

Table 1.1: Sample list of retail items offered	7
Table 1.2: Replica of budgeting calendar provided to participants	17
Table 1.3: Parameter vector for each model considered	30
Table 1.4: Sorting individuals by model	31
Table 1.5: Sorting results, percentage of participants in sample	32
Table 1s:	56
Table 3s:	56
Table 4s:	56
Table 7s:	59
Table 8s:	59
Table 2s:	61
Table 5s:	62
Table 9s:	63
Table 2.1: Comparison of group rejection of normality ... (simulations only)	93
Table 2.2: Normality tests by number of bets and average wager types	95
Table 2.A.1: Simulated 95% critical values for adjusted Jarque-Bera test	104
Table 2.B.1: Normality tests by number of bets and average wager types (all)	105
Table 2.C.1: Convergence of cumulative winnings	111
Table 3.1: Summary statistics.....	125
Table 3.2: Contract length (dependent variable), full sample.....	128
Table 3.3: Contract length (dependent variable), customer fixed effects.....	131
Table 3.4: Likelihood of cancellation/default.....	132
Table 3.B.2: Contract length (dependent variable), full sample.....	138

Table 3.B.3: Contract length (dependent variable), customer fixed effects.....	139
Table 3.B.4: Likelihood of cancellation/default.....	140

ACKNOWLEDGEMENTS

I wish to thank my dissertation advisors Vincent P. Crawford and Julie Berry Cullen for their encouragement, guidance and support. I am especially grateful to Vincent Crawford for frequent discussions throughout the development of these research papers, as well as for kindly sharing his knowledge and time in teaching me about the field of behavioral economics. I benefited immensely from conversations with Julie Cullen, whose insights on empirical issues and professional advice have been invaluable. I thank Joel Sobel for his encouragement, and for providing many thoughtful comments which improved my papers. In addition, I thank my committee members Uri Gneezy, Craig McIntosh, and Uri Simonsohn for their comments and support.

This dissertation would not be possible without the help of individuals at the anonymous institutions that generously agreed to provide the data for each of the three chapters. I gratefully acknowledge their willingness to participate and their assistance in collecting the data for me.

I am indebted to my former colleagues at the Federal Reserve Bank of Chicago whose intellectual enthusiasm often served as a personal motivation – I am especially grateful to Meredith Crowley, Michael Kouparitsas and Spencer Krane, who provided frequent valuable advice. I also acknowledge financial support from my undergraduate institution, Wellesley College via the Peggy Howard Fellowship, and the moral support of my undergraduate advisor Akila Weerapana.

I thank my fellow graduate students and the economics faculty for creating an enjoyable environment during my years at UC San Diego. Mary Jane Hubbard and Rebecca Franco graciously helped me navigate the rules and administrative aspects of the program.

Finally, I wish to express gratitude to those who I met during the job market experience – and especially to Tsinghua University, where the next journey begins.

Chapter 1 has been submitted for publication. The dissertation author was the primary investigator and author of this paper.

Chapter 2 has been submitted for publication. The dissertation author was the primary investigator and author of this paper.

Chapter 3 is being prepared for submission for publication. The dissertation author was the primary investigator and author of this material.

VITA

2001	Bachelor of Arts, Wellesley College
2001 – 2003	Associate Economist, Federal Reserve Bank of Chicago
2005	Master of Arts, University of California, San Diego
2010	Doctor of Philosophy, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Essays in Behavioral Economics: Evidence from the Field

by

Jaimie Wei-hung Lien

Doctor of Philosophy in Economics

University of California, San Diego, 2010

Professor Vincent P. Crawford, Co-Chair

Professor Julie Berry Cullen, Co-Chair

This dissertation analyzes economic behavior in three different field settings. Chapter 1 examines how students in a university dining program allocated spending of their meal points over time, given a known expiration date, fixed prices, and uniform initial endowment of meal points. I find aggregate retail sales nearly tripled their volume during the two weeks prior to the expiration date. This suggests that a more efficient spending plan may have been available for many students ex-ante. Analysis of individual

student spending profiles shows a sizable fraction of students learned how to spend their meal points more smoothly towards the end of the year. The results support existing evidence from lifecycle savings and consumption literatures that individuals often delay paying attention to intertemporal planning problems.

Chapter 2 examines the quitting decisions of slot machine gamblers using a casino dataset, asking whether the relationship between quitting decisions and winnings can be most easily reconciled using neoclassical expected utility theory or a Prospect Theory value function. Quitting decisions were highly concentrated near the monetary break-even point, consistent with Prospect Theory. Disproportionate quitting at a reference point persists conditional upon average wager size and number of bets taken, as measured by excess kurtosis. In addition, winnings distributions are positively skewed – implying that people are more likely to quit with winnings above the reference point than below it, and providing support for diminishing sensitivity of the value function’s gains and loss segments.

Chapter 3 analyzes contract choices of customers of an online layaway contractor. Participants paid a non-refundable price premium for monthly payments to be automatically deducted from their checking accounts towards the purchase of a retail item. Unlike traditional layaway, neither the price nor the availability of the item ordered was guaranteed, suggesting a key motivation for participating was as a savings commitment device. Customers selected greater monthly savings commitments for entertainment and leisure items, compared to more necessary household items. The result is robust to

customer-specific effects, price of the item, outstanding layaway obligations, and the expected month of delivery.

Chapter 1: Spending and Learning in a University Dining Program

1.1 Introduction

This paper uses field data from a university dining plan to study intertemporal consumption choices. I use a unique data set from a student room and board program at a large public university, in which an endowment of meal points can be used to pay for items in dining halls over the course of a year. The student meal plan allows precise observation of how people consume their stock of meal points over time in a natural field setting, using scanner data that report their expenditures by transaction. The meal plan can be naturally interpreted as a very simple intertemporal planning problem, in which prices are fixed, ‘lifetime’ is short and certain, and consumption expenditures largely consist of food, which is a daily necessity.

As a benchmark for measuring how students allocated their meal points over time, I assume a simple model where students’ utility of meal plan consumption is concave while their utility of cash items is linear. I then estimate and compare variations of the resulting prediction that in every week students will want to spread their meal point use approximately evenly over the year. I find that nearly half of students analyzed behaved consistently with the notion of consumption smoothing – meaning that in every week they spent approximately as though they planned to divide the remaining meal points evenly over the remaining weeks. The other half of the sample deviated substantially

from this strategy, the majority of them consistently under-consuming every week.¹ Out of the subgroup of students who were systematically off target in their spending relative to the simple model, over half managed to self-correct their ‘bias’ before the program had ended, switching discretely to spreading their meal point expenditures evenly. This self-correction occurred overwhelmingly in the final weeks of the program, suggesting that the presence of a hard deadline (and subsequent expiration of meal points) was a key motivating factor. Students definitively rejected the university-provided meal point schedule that was designed to help them budget their points over the year.

This study contributes to the field and experimental evidence on intertemporal consumption choices. Several studies test the permanent income hypothesis (PIH) in populations that rely on a predictable government transfer every month. Stephens (2003) and Mastrobuoni and Weinberg (forthcoming) show that for low-wealth social security recipients, food consumption declines over monthly increments and is highest just after receiving their social security check. Shapiro (2005) argues that the decline in caloric intake for food stamp recipients in monthly increments supports a hyperbolic discounting model. Brown, Chua and Camerer (2009) implement a savings/consumption experiment, finding support for both bounded rationality and self-control problems as an explanation for under-saving.² To my knowledge this is the first formal field study of an institutional dining program as a consumption/savings quasi-natural experiment.

¹ The analogous but opposite pattern of over-consumption may exist as well but since my data do not allow precise observation of this effect, I concentrate on students whose choices I can observe clearly, who are typically under-consumers.

² Since most of the deviations from smoothing I observe in my data imply a negative discount rate, and the motive for smoothing is not as easily established as other experimental or field studies in the

The meal plan setting has several benefits and drawbacks compared to existing experiments and field studies. First, the meal plan environment is less sterile than an experimental setting but more sterile than the ‘real world’ of government transfers. The meal plan uses a population similar to those typically used in experiments (students), but a generally different population with different objectives than those receiving social security or welfare payments. Unlike the laboratory consumption experiments, the solution (assuming the objective function is properly specified) is computationally very simple. Therefore the results are less informative about the ability to solve complex dynamic problems as emphasized in Brown et al., but more informative about the natural tendency to view everyday problems in a dynamic mindset. A drawback is that the observed data are total expenditures which in some cases include non-food items. As a result the meal plan only provides a partial picture of any single category of consumption, and does not permit any conclusions about specific categories of goods such as food or more durable items.

Since it can be difficult to infer participants’ motivations and objectives from expenditure patterns alone, I supplement the empirical analysis with results from a survey of students participating in the meal plan.³ Survey responses regarding how students intended to use their meal points show that students had one of two dominant intentions: to smooth their consumption as evenly as possible throughout the year, or alternatively that they did not think about the future when deciding how much to spend, and hence did

consumption/savings literature, I focus on the bounded rationality (learning) rather than the self-control hypothesis for dynamic planning.

³ Although the survey was conducted on students who participated in the meal plan in 2007-2008, the empirical work utilizes the 2005-2006 data. Other than an increase in the initial meal point allotment, the overall structure of the program remained identical in both years.

not see it as an intertemporal problem. Together, the survey and empirical results suggest that a delayed reaction in framing the allocation of meal points as an intertemporal task was a major reason why smoothing behavior was not as prevalent as it might have been otherwise.

To preview the results, I find that about 47% of all participants in my sample were best explained by complete consumption smoothing compared to the other models under consideration. Another 20% of participants tended to adjust their consumption in the correct direction on a week by week basis, however on average fell short of full adjustment by an estimated constant fraction. In other words, a conscious effort appears to have been made by these individuals to achieve smoothing, although the adjustment was systematically insufficient. The remaining 33% of participants were best explained by a simple form of learning, in which they started out the year consuming according to biased smoothing, but eventually adopted complete smoothing. In many cases this corresponded to a discrete and often large jump in personal spending levels. The most common week in which these students began to adopt the complete smoothing solution was in the second to final week of the school year, showing that the smoothing motive was present but realized very late into the program.

The paper proceeds as follows: Section 1.2 describes the details of the meal plan. Section 1.3 discusses the smoothing motive and other concerns. Section 1.4 describes the models under consideration. Section 1.5 describes the data. Section 1.6 describes the empirical strategy. Finally, Section 1.7 discusses the estimation results, and Section 1.8 concludes and discusses possible policy implications.

1.2 The Meal Plan

Participants in the meal plan are students living on-campus at a large public university. The plan is part of a room and board contract, so that all students who live on-campus also receive 1800 meal points. Meal points can be used at any of about ten major dining facilities around the campus, which are located directly next to each of the major residence halls. One meal point can be thought of as equal in value to a dollar, since anyone who is not enrolled in the meal plan must pay the posted prices in dollars. Meal points are stored electronically on the student's ID card, which functions as a debit card in keeping track of participants' expenditures and remaining balance. Many universities operate meal plans similar to the one described here, typically varying by frequency of expiration dates and meal point allotments.⁴

A crucial feature of the plan is that all meal points expire at the end of the year; there are no refunds or carrying over of unused points into the next year. This expiration date implies that points should never go unused unless a particular student does not have a strictly positive marginal utility of using them. At the beginning of the year, the university distributes a calendar of suggested meal point balances by week, which corresponds to complete smoothing from beginning-to-end of the initial allotment of meal points, without errors. A copy of the calendar provided to students at the beginning

⁴The rules and regulations of college meal plans are actually quite diverse, and can often be significantly more complicated than the meal plan in this paper. The trend in university dining appears to be towards offering a menu of contracts of varying amounts to incoming students, rather than offering all students an identical relatively low cost base contract, with the option to purchase more in small increments. Another common feature is for universities to offer meal plans priced by the meal or swipe (buffet style) or some combination of this and a la carte meals, whereas this particular meal plan consists only of a la carte items. Finally, some schools allow students to transfer meal points amongst themselves (often for a transactions fee), whereas this meal plan does not allow for formal transfers between accounts.

of the year is shown in Section 5.1. For different individuals, complete smoothing of the provided meal points corresponds to a different number of meals purchased at dining facilities per week. For those who consume fairly little per meal, the allotment may be enough to cover nearly all meals, while for those who consume a lot per meal, this allotment may only cover half of meals.⁵ Students can check their meal point balance easily online or by simply taking note of the balance displayed on the cash register after they have made any purchase. At the end of the year, signs are also posted in the dining facilities reminding students to finish using their meal points.

Unlike prices in the external economy, prices within the meal plan remain constant over the year. The initial 1800 meal points provided guarantees that all participants have at least this minimum number of meal points at their disposal, reducing income uncertainty effects. In addition, students' lifestyles largely consist of the predictable activities of going to classes, working, studying and socializing, the regularity of which should help facilitate smoothing. On the other hand, the year-long horizon and immediate access to the entire stock of meal points, allows ample freedom for students to either follow suggested guidelines fairly closely or else deviate from them drastically.

There is a dining facility located directly adjacent to each residence hall complex, making the meal plan the most convenient food source for students living on-campus. In addition, the meal points can be used at any of the university-run dining facilities so that if a student happens to be in a different part of campus at a given meal time, he or she can also eat in a different dining hall. Participants can rely on the dining halls to have a basic

⁵If desired, students may purchase additional meal points anytime at a discount bulk rate. Specifically, a 20 point bonus for every discrete 100 points purchased. The alternative is to pay cash and forgo the discount. In this paper I focus only on the students who did not purchase additional meal points.

menu of daily food items such as pastas, burgers, sandwiches and so on, with some daily and weekly variety in choices. Different dining facilities also specialize in slightly different cuisine types, as well as selling pre-packaged foods usually consisting of standard brand-name snacks.

The campus convenience store is another place where students can use their meal points on a greater variety of familiar non-perishables and personal items. In addition to these options, the university also runs its own online grocery store through which students can use their meal points to purchase non-perishable foods and beverages in bulk quantities, as well as cooking and school supplies. Table 1.1 below shows a partial list of items sold at the convenience and online grocery stores, but the items in the list are representative of the types of items sold through the meal plan.⁶ In summary, there should be no shortage of options for students interested in making purchases with their meal points.⁷

Table 1.1: Sample list of retail items offered⁸

convenience store:

Yoplait yogurt	Betty Crocker fudge brownie mix	Bounty paper towels
Naked juice	Aunt Jemima pancake mix	Ultra Downy
Oscar Meyer deli meats	Hershey's chocolate syrup	Glad storage zipper bags
20 fl-oz soda bottles (various)	Skippy peanut butter	Lysol
Campbell's Soup at Hand	Pepperidge Farm Goldfish	Softsoap
Easy Mac	Wheat Thins (low sodium, multigrain, ranch)	Bic razors

⁶Prices for the convenience store are about the same or slightly higher than prices for identical items off-campus. Prices for the online grocery store are shown in the table. These retail lists and prices are for the 2006-2007 school year although items and prices did not change substantially since the 2005-2006 year.

⁷A valid concern is whether all students know about all of the opportunities available where they can use their meal points. I am unable to verify this, but assume that this information is given to them during orientation or through word of mouth.

⁸ Full list of retail items available on request.

Table 1.1: Sample list of retail items offered, Continued

Poptarts	Loaf of bread	Stainless steel silverware
Honey Nut Cheerios	Kashi TLC cookies	Scotch tape
Ben and Jerry's ice cream (various)	Oreo	Shout stain remover
Cup of Noodles	Spam	Colgate toothpaste (various)
Cheez-Its (various)	Doritos	Listerine
Arizona Tea	Apple Chips	Tide detergent
Milk (vitamin D, 2%, fat free, chocolate)	Hunt's snack pack pudding	Kleenex boxes
Progresso Soup	DeCecco pasta	Paper plates and cups
Eggs	Prego pasta sauce	Alba botanica shampoo and conditioner
Tropicana orange juice, 64 fl-oz	Ritz crackers	Bicycle playing cards

online grocery store:

product name	specifications	price (in meal points)
Batteries Duracell AA	AA Alkaline Batteries, 36 pk.	27.89
Tylenol Extra Strength Gelcaps	300ct 500mg Rapid Release Gelcaps, pain reliever, fever reducer	28.99
Lotion Vaseline Intensive Care	Vaseline Intensive Care Lotion, 3/24.5 oz	15.89
Soap Bar Irish Spring	With Aloe Moisturizers, 16/4.5 oz. bars	10.69
Clorox Disinfectant Wipes	Disinfectant Wipes, 4/32 ct.	11.39
Power Bar Variety	Choc, Vanilla Crisp, Peanut Btr, 24 ct.	37.89
Shampoo Pantene	Sheer Volume, 33.9 oz.	11.59
Razor Blades Gillette Mach 3	Mach 3 Refill Razor Blades, 20 ct.	42.89
Dental Floss Glide	Glide Dental Floss, 4/50 meters	16.69
Snickers Candy Bars	Chocolate, Peanuts, Caramel, Nougat, 48 ct.	27.89
Red Bull Energy Drinks	Taurine, Caffeine, and Vitamins, 24/8.3 cans.	56.29
Starbucks Variety Frappuccinos	Caramel, Coffee, Mocha Decaf, Vanilla, 12/9.5 oz.	20.29
Martinellis Apple Juice	24ct 10oz Bottles Pure Apple Juice	28.99
Arrowhead Mountain Spring Water	Sport Top, 28/24 oz. Bottles	12.89
Diet Coke	24pk Cans	11.99
Rice Calrose White	Calrose White Rice, 50 lb. Bag	21.99
Paper Hewlett-Packard Printer	500ct, HP Bright White, 24lb Inkjet Printer Paper	15.89
Notebooks	Tradition Three-Subject Spiral Notebooks, 6pk, 120 Sheets each	11.49

Table 1.1: Sample list of retail items offered, Continued

Post-it Notes	Assorted Colors, 3 x 3 inches, Larger Pack 2100 sheets 21/100 pk.	18.49
Printer Ink Refill System	480ml Premium Inkjet & Photo Ink for most printers: black, cyan, magenta, yellow	37.99
Pens Uniball Rolling	Fine Point, Blue Ink, 12 ct.	10.99
Leisure Beach Mat Jungle Hibiscus	Tiki 3ft x 6ft Leisure Polypropylene Mat	14.99
Rice Cooker Aroma	Rice Cooker and Food Steamer, 10 cup size, Keep Warm Function, 15 Hour Delay Timer, Digital Control Panel	53.49
Indoor Grill	Oster Indoor Grill, 170 sq Inch Non Stick Grilling Surface, Drip Tray, Dishwasher-safe Base and Grill. Limited Time Only.	59.99

The main sources of competition for the dining facilities are several on-campus fast food style establishments and shopping for groceries off-campus and cooking in the residence hall kitchens. Some students may in fact prefer these alternatives, and almost all students will have to supplement their meal plan with some of these outside options in order to satisfy their overall food demand, since 1800 dollars for the school year only amounts to \$7.50 a day.⁹

There is no way to give an objective assessment of the quality of food at this particular institution.¹⁰ However, a key point is that meal points can be used for purchasing more than just prepared foods, as shown in the preceding list. The fact that these items are brand-name sets a known lower-bound on their quality, and greatly reduces the concern that some students are incapable of spending 1800 meal points over the course of the year.

⁹ Calculation includes weekends but not school breaks.

¹⁰ A website that posts user reviews of restaurants gave a 4 out of 5 star rating (N = 13) to the largest of the institution's campus dining facilities. Selection problems are a concern. See Dellarocas (2003) for a survey of online reviews.

1.3 Discussion of smoothing motives

An immediate concern is whether or not it is valid to conclude that smoothing meal point use is the optimal solution when the reality is that students can eat their meals anywhere. This conclusion arises directly out of the assumption that utility is concave in any items purchased with meal points, while utility is linear in all other items – seemingly a reasonable assumption since meal points can only purchase a limited set of goods. If the proposed utility function is incorrectly specified however, there is nothing inherently suboptimal about a student spending all of his meal points in one week – similar to the idea that utility functions exist that make saving almost nothing for retirement optimal. Therefore, I use the proposed model not necessarily to assess the optimality or sub-optimality of students' decisions, but rather as an intuitive benchmark that can be used as a comparison point for actual behavior. In fact, I find that nearly 80 percent of all program participants were statistically characterized by some degree of smoothing behavior – although a sizeable fraction of them only began smoothing their spending late in the year, often resulting in a discrete jump in expenditure levels. For this kind of pattern to be fully rational in the backward induction sense of dynamic planning, those students would have had to anticipate from week 1 that they would have a major shift in utility for meal point items late in the year. Although this explanation does not seem likely, particularly if the shift in utility for most people coincides with the impending end of the program, I cannot definitively rule it out. Throughout the paper I will use the word 'optimal' as it relates to the solution of the proposed model, and generally withhold discussions about welfare implications.

The smoothing motive earns some support through the survey evidence (see Appendix for details), in which the most frequently (42 percent) desired shape of meal point use was to smooth evenly over the year.¹¹ For first-time participants in the meal plan, the proportion who wanted to spread point use evenly, is even higher at 52 percent. As one survey respondent explained, “(There is) no point saving any in the end because the non-food stuff they sell at the end of the year is overpriced, and you don’t get extra meal points back (refunded).” Another respondent said, “(I) need (my) meal points to last the whole year, that way I don’t need to spend more money.”

I use quasi-linear utility to describe students’ preferences over items they can purchase with their meal points and all other items that can be purchased using cash. Given the structure of the meal plan, quasi-linearity is appealing because the use of meal points to purchase items is limited in both scope of goods that can be purchased and the time frame of use – while in contrast cash on hand can be used to purchase anything at any time. It is reasonable to suppose that there is diminishing marginal utility of consumption for the fixed subset of goods offered within the meal plan. For non-meal plan items however, the universe of items is boundless. I assume that the student can choose items outside the plan in a way that makes him marginally better off dollar for dollar, rather than in a diminishing manner as in the confines of the meal plan.

1.3.1 Censoring and selection effects

In order to simplify the model and analysis, I target those students whose optimization problem is as close as possible to a simple ‘cake-eating problem’ with a

¹¹The overall most frequently chosen answer to this question was to use meal points ‘as necessary’ (44 percent), which does not necessarily correspond to a particular shape.

cake of size 1800 meal points. This means that I exclude any students who purchased additional meal points at some point during the year, as well as those who simply finished their meal points before the year had ended. Students who purchased additional meal points typically also finished their original meal points early, although those who finished early did not always choose to purchase additional meal points.¹² The purpose is to capture to the greatest extent possible, only the subset of the population for whom the simple model set up of the planning problem is likely to apply. The model does not address using cash to purchase meal plan goods (as someone who finished their meal points early might opt for), and also ignores the decision to purchase additional meal points.

Censoring the sample to exclude students who purchased additional points as well as anyone who finished their initial allotment early, results in selecting students with a lower meal point demand relative to their excluded peers. I do not try to assess whether the excluded students spent and purchased additional meal points following any particular smoothing or permanent income model.¹³ Another motive for focusing only on those who consumed only 1800 meal points is that it helps exclude those for whom uncertainty might have existed (through income or otherwise) about whether or not they eventually would purchase additional meal points. If we imagine a scenario where meal points are not pre-paid, the individuals that remain in the sample should have a meal point demand either at or below 1800 points. This places even greater importance on the expiration of meal points, since the consumption behavior of many students in this

¹² In the data section I describe in detail how I filter out those students who finished their meal plan early.

¹³ The fact that meal plan items can also be purchased with cash makes motivations on additional point purchase/usage difficult to disentangle with the observed data.

sample will be primarily driven by an effort to avoid letting their meal points expire. According to program administrators, a common complaint by students and parents is that 1800 meal points is not enough, and high-demand students often attempt to stretch 1800 points over the entire year rather than buying more meal points.¹⁴ In addition to low-demand students, students who were successful at this self-restraining attempt would remain in the sample.

1.3.2 Budgetary and taste heterogeneity

In spite of one commonly held view that college students are budget-conscious, an opposite concern might be that students come from wealthy families, and are being provided with ample cash from their parents throughout the year that sustains high levels of food consumption outside the meal plan. This would imply a low personal valuation of meal points by students relative to their outside options. Unfortunately, I do not have access to students' individual financial information. However, 76 percent of all meal plan participants were recorded as having received some form of financial aid that year (including loans). The fact that such a high percentage of meal plan participants were receiving financial aid refutes the idea that meal points were of low intrinsic value. However, the concern that meal points had low valuation relative to cash items may very well have been valid for a subset of participants. The Appendix includes a summary of survey responses to specific questions designed to gauge the importance of this and related concerns.

¹⁴ Several students in the survey also mentioned being motivated by not wanting to purchase any additional meal points.

Substantial heterogeneity in cash resources is likely to exist among students, which in turn influences the relative importance of following the proposed optimal plan. That is, certain students are much more likely to be more able to ‘afford’ deviations from optimal smoothing than others, or put another way, the meal plan itself is of varying importance to different students. This feature is left out of the model due to data limitations (mainly not knowing how much cash an individual student has), although it is indeed a practical shortcoming in the analysis.

A related point concerns the possible inferiority of meal plan goods in the hypothetical case where no meal plan contract existed. If a student’s total wealth (including cash) declines steadily over the year, it is possible that the demand for meal plan items in turn increases via their inferiority by definition of an inferior good. If this is the case, then any increases we might observe over time in students’ meal plan use, are confounded with the fact that the meal plan is a finite horizon problem. While inferiority of the set of meal plan items could be a reality for some students, the timeline of expiration of meal points is a reality for all students, it seems reasonable to believe that the expiration date of meal points and their cost being already sunk is the overriding incentive, and not that a decline in total wealth makes meal plan items seem more appealing. This would especially be the case if increases in meal plan use are large and discrete in nature as they are in the data, rather than gradual.

Another concern is that students are variety-seeking in their food consumption, and so might not be expected to smooth meal point usage over short intervals of time. Over a large enough interval of time however, we would expect that variety-seeking

averages out so that the smoothing prediction can be realized empirically. In the case of the meal plan, a reasonable candidate for that time interval is weekly, since classes run on a weekly schedule. The data show that over half of students have a distinct trend in their meal plan usage that is generally not reconcilable with variety-seeking, which tends to predict noise when not aggregated enough over time.

1.3.3 Meal point trading

A final point to note about the structure of the plan is that since meal points are stored on each student's university ID card, transferring or selling points to other students requires additional coordination effort. Survey responses indicated that giving away meal points to friends or roommates was common when a student found themselves with more meal points than desired, but selling them does not seem to be prevalent. Since a student's ID card is needed in order to utilize other university facilities including the gym, libraries, shuttle buses, and occasionally to verify students' identity during exams, it would not be convenient for students to trade ID cards on a long-term basis solely for meal plan purpose

1.4 Proposed models

I test four model alternatives that describe how students might approach using their meal points, and compare their explanatory abilities for each student. The goal is to verify the extent to which students followed either the school's guidelines, or the extent to which they smoothed meal point consumption. I estimate the following models using maximum likelihood and compare their explanatory abilities statistically: 1. The school's

recommended plan, 2. Complete smoothing, 3. Biased smoothing, and 4. Biased smoothing with learning. The subsections below describe each of the models.

1.4.1 School's recommended plan

The predictions of the school's suggested guidelines are self-explanatory by looking at Table 1.2 below. Students are given a copy of this schedule for personal reference, but are not provided with any external reward or penalty for following it. This suggested plan corresponds to the very first week's complete smoothing plan for the entire year, and does not account for individual students' previous weekly errors as the year goes on. In other words, given a previous week's errors, a student trying to follow the school's suggested guidelines will generally have to make a more drastic current week correction in order to be on schedule with the school's plan.

In comparison, this is a fairly strict criterion to follow throughout the year, although it is feasible that some students might make a deliberate effort to spend on schedule by following this recommended plan. If a student goes the whole year without making any errors, the school's suggested plan and the complete smoothing solution would fit equally well.

Table 1.2: Replica of budgeting calendar provided to participants

Based on 53 meal points per week. Meal points expire at the end of spring quarter, are non-refundable and non-transferable.

<u>Fall quarter</u>		<u>Winter quarter</u>		<u>Spring quarter</u>	
<i>orientation</i>	1800				
week 1: Sunday, Sept 25:	1747	week 1: Sunday, Jan 8:	1164	week 1: Sunday, April 2:	581
week 2: Sunday, Oct 2:	1694	week 2: Sunday, Jan 15:	1111	week 2: Sunday, April 9:	528
week 3: Sunday, Oct 9:	1641	week 3: Sunday, Jan 22:	1058	week 3: Sunday, April 16:	475
week 4: Sunday, Oct 16:	1588	week 4: Sunday, Jan 29:	1005	week 4: Sunday, April 23:	422
week 5: Sunday, Oct 23:	1535	week 5: Sunday, Feb 5:	952	week 5: Sunday, April 30:	369
week 6: Sunday, Oct 30:	1482	week 6: Sunday, Feb 12:	899	week 6: Sunday, May 7:	316
week 7: Sunday, Nov 6:	1429	week 7: Sunday, Feb 19:	846	week 7: Sunday, May 14:	263
week 8: Sunday, Nov 13:	1376	week 8: Sunday, Feb 26:	793	week 8: Sunday, May 21:	210
week 9: Sunday, Nov 20:	1323	week 9: Sunday, Mar 5:	740	week 9: Sunday, May 28:	157
week 10: Sunday, Nov 27:	1270	week 10: Sunday, Mar 12:	687	week 10: Sunday, June 4:	104
Finals week: Sunday, Dec 4:	1217	Finals week: Sunday, Mar 19:	634	Finals week: Sunday, June 11:	51
<i>end of quarter total</i>	1164	<i>end of quarter total</i>	581	<i>end of quarter total</i>	0

1.4.2 Complete smoothing

The predicted consumption under the complete smoothing plan comes from solving the optimization problem below in every week of the meal plan. Weekly utility is given by $U_t = u(C_{n,t}) + a_{n,t}$ where $C_{n,t}$ denotes meal points used in week t by student n , and $a_{n,t}$ denotes cash used in t by student n . $W_{n,t}$ denotes the stock of meal points available to student n at the beginning of week t .

$$\text{Max}_{\{C_{n,j}, a_{n,j}\}_{j=t}^T} u(C_{n,t}) + a_{n,t} + \sum_{i=t+1}^T u(C_{n,i}) + a_{n,i}$$

subject to

$$W_{n,t} = W_{n,t-1} - C_{n,t-1}$$

and

$$W_{n,t} \geq 0, C_{n,t} \geq 0$$

and $W_{n,1}$ is given. I assume there is no time discounting since the primary good in question is food, which has a strong smoothing motive. In addition, survey responses suggest that standard time-discounting did not play a large role, given that almost none of the respondents wanted to use more points earlier in the year and fewer points later in the year.¹⁵ So long as $u(C_{n,t})$ is monotonic and strictly concave in $C_{n,t}$, the solution will equate marginal utilities in the current and future periods, giving complete smoothing going from week t forward as the week t solution. The optimal smoothing plan in week t can be expressed as the following fraction of the current stock of meal points. Throughout the paper, $C_{n,t}^*$ will denote this complete smoothing solution.

$$C_{n,t}^* = W_{n,t}/(T-t+1)$$

In other words, the prediction about behavior is that every week the student will re-optimize, which amounts to simply dividing the remaining meal points evenly by the number of remaining weeks. They will plan to consume that same amount each week into the future as part of the solution, but in fact the very next week they will revise the plan. In other words, the student might make ex-post mistakes, but will correct them by attempting to hit the new consumption target in the next week $C_{n,t+1}^*$.

To make sure the implications of complete smoothing are fully clear, here is a numerical example of how this plan might be implemented. Suppose that in the first week a student's plan under optimal smoothing dictates consuming 54.55 meal points (1800 points/33 weeks), leaving him with a balance of 1745.45 meal points remaining.

¹⁵In contrast, Shapiro (2005) and Mastrobuoni and Weinberg (2008, forthcoming) focus on individuals' discount rates and in particular of the quasi-hyperbolic form, which is more plausible in the settings they consider, since they find that individuals generally don't have enough money for food. Here a significant fraction of students appear to have the opposite problem.

However, suppose in practice he ended up consuming 80 meal points in that first week, leaving him instead with a balance of 1720. In the second week he would aim to consume 1720 meal points divided by 32 weeks, or 53.75 meal points a week. In other words, for his slight overconsumption in the first week, he must eat very slightly less in all the subsequent weeks. Notice that mistakes that happen early in the year impose a smaller change in the student's weekly plan going forward, than those that occur late in the year. For an example of the latter, suppose a student had been following the recommended plan on average, up until the 31st week. This means that at the beginning of the 31st week she has 163.65 meal points left to smooth over the next 3 weeks, at the recommended 54.55 meal points per week. Suppose she over-consumes in the 31st week by using 80 points instead. This leaves her with 83.65 points to spread over the remaining two weeks in the plan, or just 41.83 per week. The fact that early mistakes do not count as much as later mistakes, gives students some degree of leeway to learn the smoothing solution without excessive penalty.

1.4.3 Biased smoothing

Now we also allow for the possibility that the student might be consistently 'off target' or biased, in his weekly adjustment toward the targeted $C_{n,t}^*$. Such bias in the meal plan context could arise from a few possible sources in the context of the previous model. The most obvious candidates are consistent misjudging of either of two variables in the problem that are required to implement the smoothing solution: the number of weeks remaining before points expire, and the number of meal points remaining in the student's account. This type of data does not allow us to distinguish which one (or both)

of these two variables is the source of the spending being off-target. Therefore the goal is to simply quantify the degree to which a simple bias in smoothing as described in the previous section determines the realized consumption plan, without attempting to nail down the specific source of the bias.

Camerer and Chua (2009) report that their experimental subjects tend to deviate from optimal dynamic plans by consuming a constant fraction of the optimal plan. They find that this explained behavior better than so called rule-of-thumb behavior, in which subjects were predicted to consume a fixed fraction of their total income or cash. Consuming a fixed fraction of total income or cash, makes some sense as a rule of thumb when income payments are continually coming in. However in the meal plan setting, consuming a fixed fraction of total meal points is not even a close approximation to an optimal plan since all of the meal points are provided up front. Consuming a fixed fraction of the optimal smoothing plan on the other hand is more plausible, and so I compare that possibility to the case of complete smoothing itself.

Allowing for bias in weeks $t = 1$ to $T-1$, we have

$$C_{n,t,\text{biased-smoothing}} = \lambda_n C_{n,t}^* = \lambda_n W_{n,t} / (T-t+1)$$

where λ_n is the student-specific fraction of the perfect-smoothing consumption level actually consumed, and may be either greater than or less than 1. In the final week T , I assume

$$C_{n,T,\text{biased-smoothing}} = C_{n,T}^*$$

since it is not possible for the student to over-consume beyond his resource of points, and under-consuming in the final week predicts forgoing resources. Note that complete smoothing is the special case where $\lambda_n = 1$.

1.4.4 Biased smoothing with learning

The next model is a hybrid of the previous two, in which students may optimize with bias initially but eventually catch on to the complete smoothing solution. One rationale is that there may be less urgency to optimize too carefully at the beginning of the year since they still have the entire school year ahead of them. Another rationale is that it may take time or some external trigger to learn how to smooth accurately.

I consider a very blunt measure of learning, in which at some week in the program the student makes a one-time switch from biased smoothing to complete smoothing. I interpret this as the student's conscious realization of the entire intertemporal problem, whereas he or she may not have fully realized the problem before. Assuming that students smooth in a biased fashion initially gives them a substantial benefit of the doubt with respect to rationality. Realistically, students may be consuming based on a flat line or another arbitrary method, but I assume that there is at least some degree of tracking the complete smoothing path however biased or imprecisely estimated it might be. The predicted consumption paths are therefore given by

$$C_{n,t,\text{learning}} = \begin{cases} C_{n,t,\text{biased-smoothing}} = \lambda_n \cdot \frac{W_{n,t}}{T-t+1} & \text{for } t \in [1, k_n - 1] \\ C_{n,t}^* = \frac{W_{n,t}}{T-t+1} & \text{for } t \in [k_n, T] \end{cases}$$

Note that when $\lambda_n = 1$ the student's entire consumption path can be characterized by complete smoothing, hence no students in this category can have $\lambda_n = 1$.

1.5 Data and preliminary analysis

1.5.1 Censoring the sample

The sample consists of all meal plan participants in the 2005-2006 academic year who did not purchase more meal points at any point in the year, and also did not finish the initial allotment early.

As mentioned earlier, I exclude those who purchased additional points to eliminate possible uncertainties about future resources. This also eliminates the possibility that non-optimal point purchase behaviors confound the appearance of consumption smoothing.¹⁶ In addition, to address the concern that students who run out of meal points early may decide to eat in the exact same dining hall locations but pay cash instead, I also disregard all students who finished their meal points early by any number of weeks.

I define finishing early as in any week having less than 10 meal points in the account and at the same time spending 0 meal points. This is meant to be a proxy for running out, or having too few meal points left to find it worth spending. Note that skipping dining hall food with a under 10 point balance for one week, then clearing the

¹⁶ Less than perfectly rational point purchase behavior can also interfere with the perception of smoothing meal point usage. Take for example a student who after using the initial 1800 points, takes a two week break from his meal point account by paying cash instead, and only after the two weeks is over realizes that he or she eats enough in the dining halls to warrant purchasing an extra 120 points (he would have purchased even more had he realized this earlier).

balance in a future week by spending everything, still counts as finishing early.¹⁷ After also dropping select individuals who left the school mid-year and other special cases that invalidate the planning problem, there are a total of 4471 individual consumption profiles remaining. 1839 students or 30 percent of all original participants were eliminated in total.¹⁸

I observe student expenditure levels as recorded on their debit card ID, but not the specific items they purchased. I follow much of the empirical consumption literature by equating expenditures and consumption, which tends to be reasonable for goods that are perishable. Although the scanner data from cashier stations at meal plan locations is at a real time frequency, I aggregate each individual's purchases on a weekly basis to reduce noise, variety-seeking and day-of-the-week effects. Since individual data tends to be quite noisy, estimation would not benefit from disaggregating to a higher frequency.

1.5.2 Coordination of start-date

I consider the academic year as 33 weeks long, not counting winter and spring breaks when the dining facilities are closed.¹⁹ In the data, some students arrived a few days earlier than others, so I begin analysis for everyone on the same day, September 24th 2005. This means that although everyone starts with 1800 points, some individuals will have fewer points in the estimation if they began consuming prior to that day. However,

¹⁷ The rationale is that some students may see 10 points as too few to purchase a meal, yet when they are reminded by friends or posted signs to finish using their meal points in the final week, they might remember to do so despite the low balance. I still want to exclude these individuals since the 1800 points was really not enough for them.

¹⁸ Examples include students who withdrew from school mid-year or spend one or more quarters away or at a different campus.

¹⁹ Each week is defined to contain the same number of weekdays and weekends to the extent possible (either 7 or 8 days, depending on the week). Weeks near school holidays and vacations will have reduced aggregate consumption volumes, but this does not appear to correspond with trends seen in individual profiles.

this does not affect the nature of the problem, it simply allows students who arrived early to start with a slightly lower allocation of meal points. For example, suppose a particular student arrived a few days early, spending 30 of his meal points before the designated starting date. His initial balance W_1 would just be $1800 - 30$, or 1770 meal points.

1.5.3 Preliminary data analysis

Figure 1.1.1 shows the aggregate consumption profile for the sample of participants we empirically analyze. This gives a general picture of campus-wide dining hall purchase volume for meal plan participants.²⁰ Expenditure volume tends to rise as the year progresses, with a sharp increase in the final week.

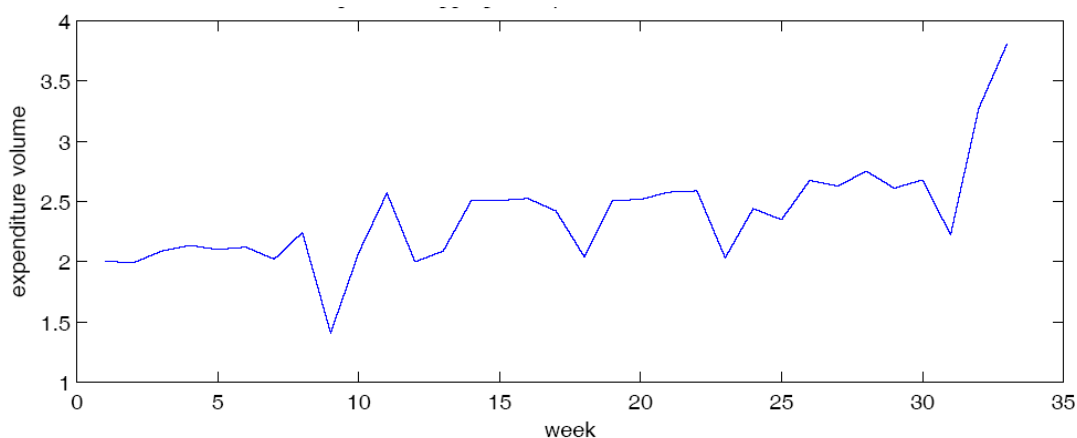


Figure 1.1.1: Aggregate expenditure volume in 10,000's of meal points, (N = 4471)

We see also from Figure 1.1.2, that by the end of the program, most participants have not wasted large sums of points, with nearly 80 percent of students letting no more than 10 meal points expire. The fact that students are not willing to forgo their meal points shows that the points do in fact have value to the students. Figure 1.1.3 shows in

²⁰Average consumption takes the same shape, since we just divide aggregate consumption by the number of individuals, a constant.

greater detail the participants who left more than 100 meal points in their accounts to expire. There are few of these individuals, totaling just 69 participants.

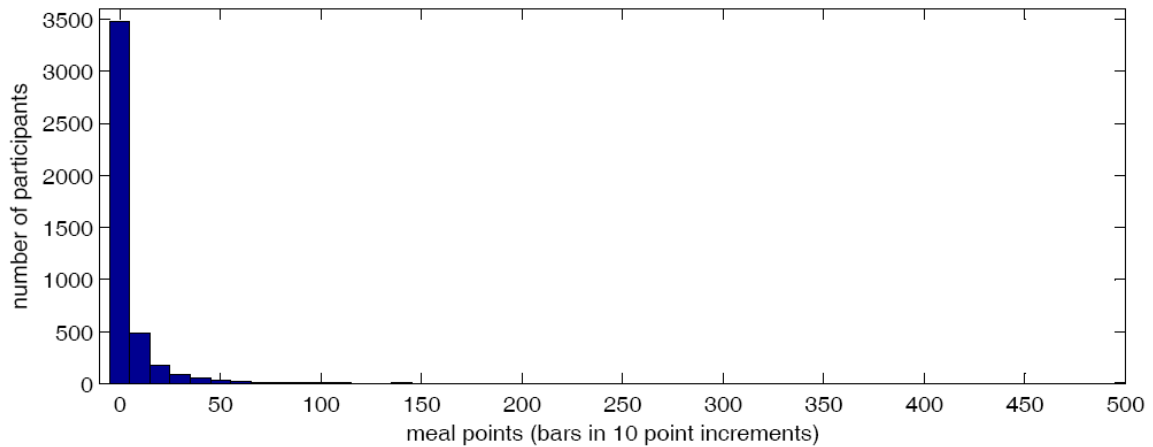


Figure 1.1.2: Remaining balances at end-of-year, (N = 4471)

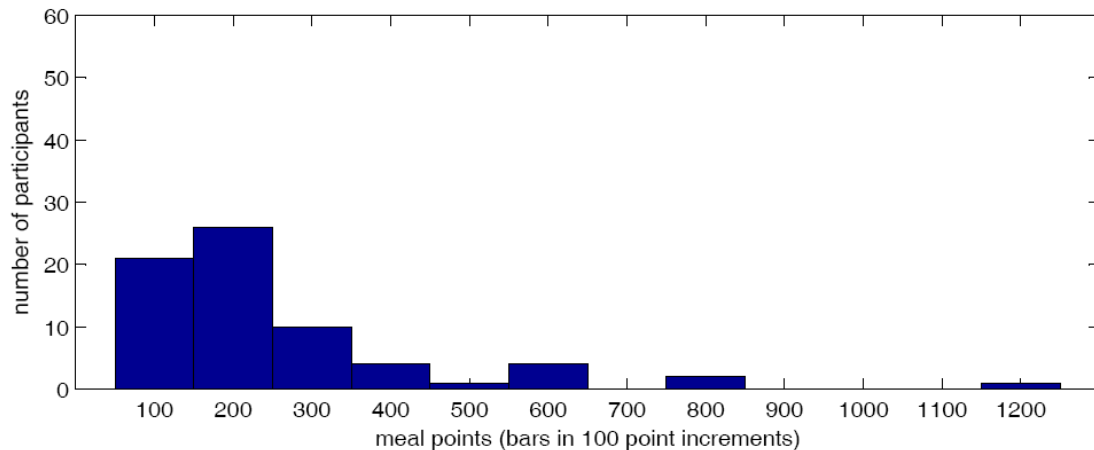


Figure 1.1.3: Remaining balances (greater than 100 only) at end-of-year, (N = 69)

It is also informative to know how often participants actually use the meal plan, both in terms of approximately how many meals per week were consumed and the average expenditure per meal. In order to gain insight regarding how spending is occurring beyond each student's aggregate activity, I group each student's expenditures into four types: breakfast, lunch, dinner, and retail, and count the number of each type of meal for each student. I define breakfast broadly as any purchase that occurred in a dining

facility between the hours of 2:00 AM and 10:59 AM. Lunch is defined as any purchase that occurred in a dining facility between the hours of 11:00 AM and 4:59 PM. Dinner is defined as any purchase that occurred in a dining facility between the hours of 5:00 PM and 1:59 AM.²¹

For each student I tabulate the number of days in each week that they make a breakfast, lunch or dinner purchase and calculate the proportion of meals eaten each week under the meal plan assuming three meals a day. I also construct an alternative measure for proportion of meals eaten under the meal plan by excluding breakfast (hence assuming only two meals a day).

²¹These are imperfect measures since some people may like to eat dinner at 4:30 PM or lunch at 10:30 AM. However, what is more important is whether they had a transaction during these blocks of time so that I can count approximately how many 'meals' were consumed by a student using the meal plan. But clearly, if someone purchased a meal at 10:50 AM and went back for a drink at 11:10 AM as part of the same meal, this would result in an over-counting of meals. Unfortunately, since I cannot tell what kind of item they purchased, this coarse measure has to suffice.

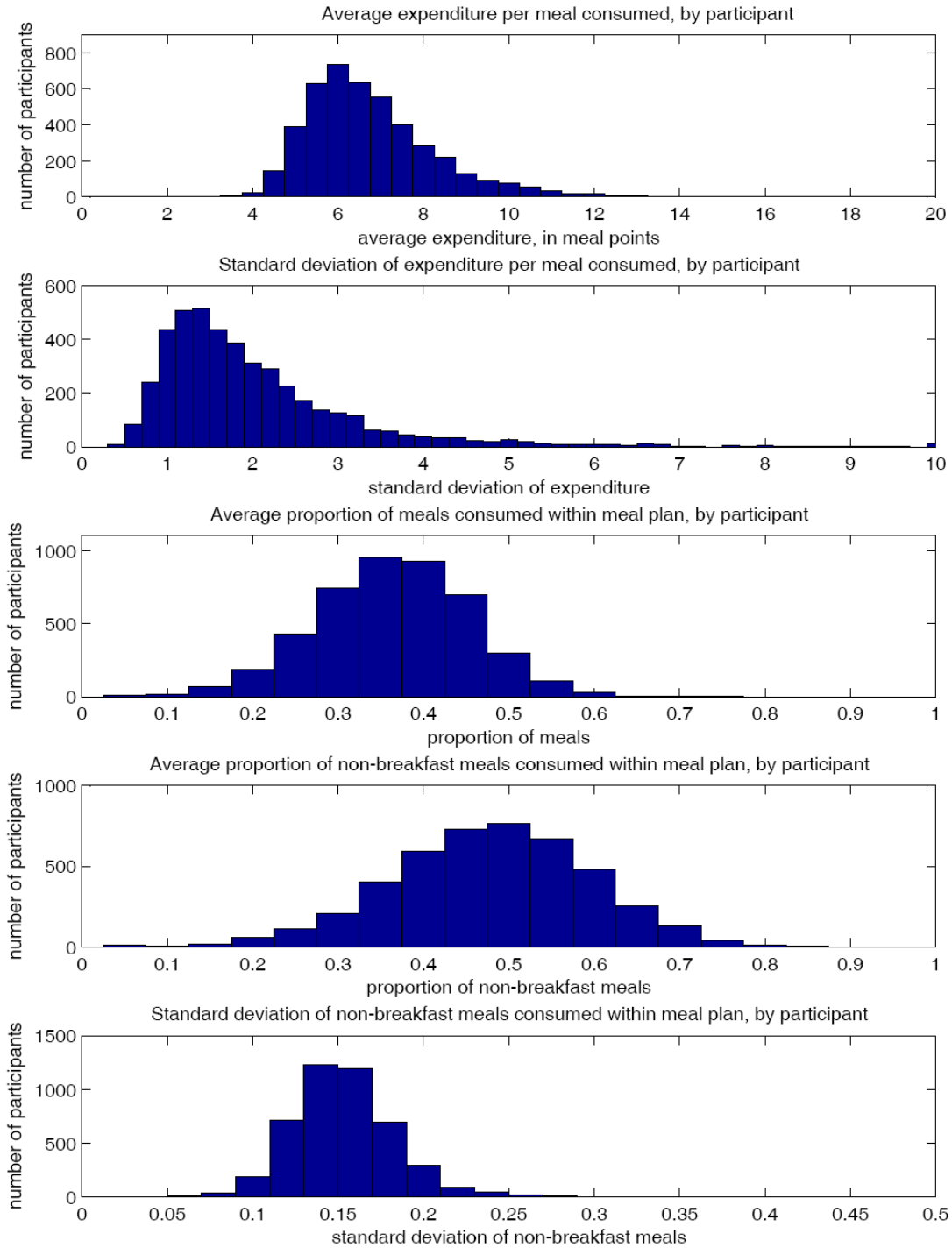


Figure 1.2: Meal distributions, by participant

Figure 1.2 shows the mean and standard deviation of these ‘proportion of meals consumed in plan’ measures. The inclusion or exclusion of breakfast changes the

distribution of meal shares within the plan substantially. The peak of the distribution is at about 35 percent of meals when breakfast is included, and 50 percent of meals when breakfast is excluded. This is not surprising, since breakfast is typically the smallest meal of the day as well as the earliest, making it most prone to being skipped.

The distributions also confirm that students are not eating every meal using the meal plan, but are supplementing with outside alternatives. The within-student standard deviation in the bottom graph, suggests that students have a reasonable amount of flexibility in deciding how many meals to eat in the dining halls and how many to eat outside each week.

I define a retail purchase as any purchase made to the online grocery store or to the convenience store/coffee shop in an amount above 20 meal points (used as an upper-bound for what qualifies as either a snack or meal – notice that the distribution of individual per meal expenditure means is below this cut off). In other words, single transactions at the convenience store that are less than 20 meal points worth are still categorized as breakfast, lunch or dinner by the time stamp criteria above. Transactions over 20 meal points made in locations other than the online grocery, convenience store/coffee shop are also still counted as meals. The purpose here is to get a measure of goods purchased by students that are more durable or bulk in nature. Figure 2m shows the aggregate volume and frequency of these retail purchases by week, with purchase frequency (per week) on the left axis and purchase volumes shown on the right axis. Both retail purchase volume and frequency rise gradually over the course of the year and very sharply in the last month of the program, more than doubling each week for the two

weeks prior to the final week. This trend is counterintuitive to the usefulness of bulk or durable items, since purchasing such items earlier in the year would yield more opportunities to use them than purchasing them late in the year.

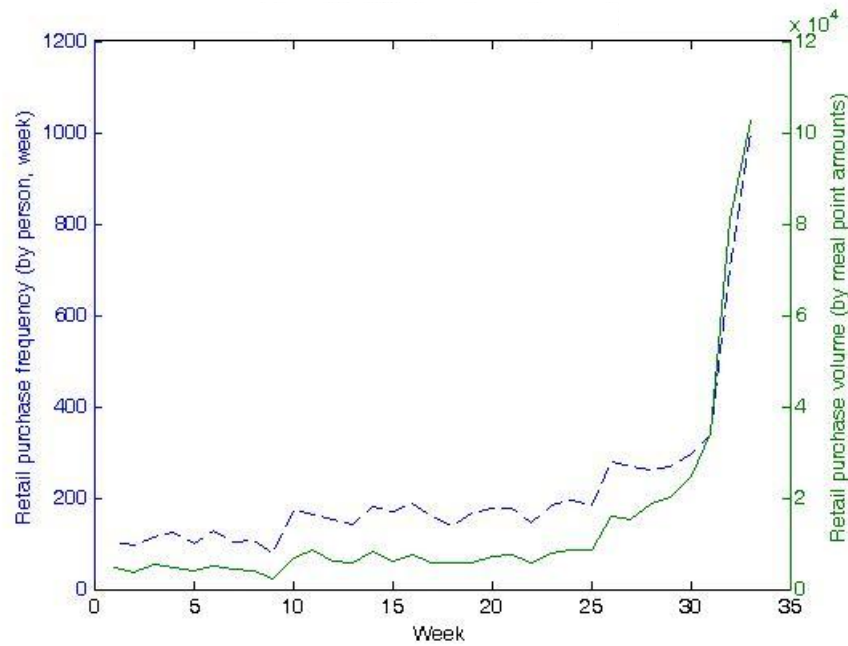


Figure 1.3: Retail purchases, by week

1.6 Estimation

I estimate each of the above models on each individual student using maximum likelihood. Individual-specific estimates of all model parameters help to address the high degree of heterogeneity in the consumption profiles. Although I do not observe the exact week that a student ‘consumes’ the item he or she has purchased, I equate expenditures to consumption, which according to Figure 2m is likely to be fairly accurate for all but the uptick in retail volume that occurs in the last three weeks. The data are assumed to follow the process

$$C_{n,t,\text{actual}} = C_{n,t,\text{model}}^* + \varepsilon_{n,t,\text{model}}$$

where ‘model’ subscript refers to either the school’s suggestion, complete smoothing, biased smoothing, or biased smoothing with learning. $E_{n,t,model}$ are in the case of each model, assumed to be independently and identically distributed $N(0, \sigma_{n,model}^2)$ in the hypothetical case of no budget constraints. However, with the imposed budget constraints, the error distribution in each week t will be truncated by $-C_{n,t,model}^*$ on the left (consume nothing), and $W_{n,t} - C_{n,t,model}^*$ on the right (consume everything). So the actual error distribution used in the estimation will be truncated normal, $\epsilon_{n,t,model} \sim N(0, \sigma_{n,model}^2, -C_{n,t,model}^*, W_{n,t} - C_{n,t,model}^*)$ to account for the individual moving endpoints of the distribution each week. This normalization will make a larger difference in the initial weeks and final weeks when the student is near the resource constraint.²²

Table 1.3 below shows the vector of parameters estimated for each model.

Table 1.3: Parameter vector for each model considered

	<i>School's suggestion</i>	<i>Complete smoothing</i>	<i>Biased smoothing</i>	<i>Biased smoothing with Learning</i>
$\sigma_n(stddev)$	yes	yes	yes	yes
$\lambda_n(bias)$		=1	yes	yes
$k_n(learningweek)$				yes

The individual’s log likelihood function is given by the sum of log probability density functions of random variable $\epsilon_{n,t,model}(\beta_n)$ over all time periods.

$$\ln L = \sum_{t=1}^T \ln \left(\frac{\left(\frac{1}{\sigma_{n,model}\sqrt{2\pi}} \right) \cdot e^{\frac{-\epsilon_{n,t,model}(\vec{\beta}_{n,model})^2}{2\sigma_{n,model}^2}}}{\Phi(-C_{n,t,model}^*(\vec{\beta}_{n,model})) - \Phi(W_{n,t} - C_{n,t,model}^*(\vec{\beta}_{n,model}))} \right)$$

where $\Phi(\cdot)$ is the cumulative distribution function of a $N(0, \sigma^2)$ random variable.

²² Goldman (2000) shows that MLE with truncated errors is asymptotically normal.

The learning model is estimated the same way except that we need to scan over all possible starting weeks in the range $k_n \in [1, T]$, choosing the one whose remaining estimated parameters best maximize that particular student's log likelihood. The estimating procedure for each student can be described as

$$\text{Max}_{k_n} \left\{ \text{Max}_{(\sigma_n, \lambda_n)} \left\{ \ln L(\sigma_n, \lambda_n, k_n) \right\} \right\}$$

Or in other words, find the optimal vector (σ_n, λ_n) first for each possible k_n , then select the week k_n which produces the best likelihood of generating the observed consumption.

1.6.1 Sorting of students by estimation results

For each meal plan participant in the sample, the goal is to find the model that fits each person best. Since we have a combination of nested models and non-nested models, I use the following procedure to find the best fitting model for each student. First, likelihood-ratio tests are used to distinguish between biased smoothing and complete smoothing. From there, I compare the result to the learning model using the Vuong (1989) test for non-nested models.

Table 1.4: Sorting individuals by model

SM = Complete smoothing, BIAS = Biased smoothing, BWL = Biased smoothing with learning			
$LR(BIAS, SM)^*$	$Vuong(BWL, BIAS)^{**}$	$Vuong(BWL, SM)^{**}$	RESULT
BIAS	BWL with $\hat{k}_n \neq T$ or ambig		BWL
BIAS	BIAS or BWL with $\hat{k}_n = T$		BIAS
SM		BWL	BWL
SM		SM or ambig	SM
*favors listed model at 95% level		**favors listed model at 90% level	

Since the Vuong test often is unable to accept or reject either model, I categorize ambiguous results from the Vuong tests into the most rational alternative, which stacks

the deck against falsely accepting any behavioral biases. In other words, if the test for the learning model and complete smoothing is ambiguous, I put the student into the complete smoothing category. If the test for the learning model and biased smoothing is ambiguous, I put the student into the learning category unless the week that learning occurred was the final week in the program since this makes learning equivalent to biased-smoothing.²³ The rationale for this is that coming to the realization of needing to finish the meal points in the final week is less about learning than simply adhering to the budget constraint. Table 1.4 above shows the combination of likelihood ratio (LR) and Vuong tests that must be passed in order for a student to be classified into each model.

Finally, from the best of the three models from the procedure just described, I again use a Vuong test to compare the result from the above table to the school's suggested spending plan, and re-categorize accordingly. The tests from this very last step overwhelmingly reject (or in a few rare cases, are unable to distinguish from) the school's plan as an accurate model of individual behavior throughout the year, resulting in overall rejections of the school's suggested plan. Students appear to adjust more sensibly to their previous spending volatilities, than the irrational rigidity of the school's suggestion. The distribution of students in each model type is displayed in Table 1.5 below.

Table 1.5: Sorting results, percentage of participants in sample

	<i>90% level*</i>	<i>95% level*</i>
Complete smoothing	46.4%	47.7%
Biased smoothing	20.6%	20.5%
Biased smoothing with Learning	33.0%	31.8%
School's suggested plan	0.0%	0.0%

*denotes significance level of Vuong tests. LR tests always at 95% level.

²³ I still allow the possibility that $k_n = T$ because otherwise the estimation routine is more likely to find a superfluous learning week against which biased-smoothing would have to compete.

1.7 Results

The maximum likelihood estimations and subsequent sorting, generates results for each individual participant in the meal plan sample. The individual nature of the analysis poses a challenge on how to display the results since there are a large number of participants. So for each model estimated, I discuss some general trends within each model and a few individual consumption profiles within each category. It is not very informative to view profiles in the aggregate, so I select a representative or random sample. Although the differences between the 90% significance levels and the 95% significance levels are slight, the discussion that follows will be for the 90% level.

In the graphs of individual consumption profiles, each student's actual data is represented by the blue line with open circles representing actual consumption levels. Solid red dots indicate where the current model predicts consumption should be, so that when a red dot fills a blue open circle, the student has hit the model's predicted consumption exactly. Dashed lines indicate the predicted path of other relevant models for comparison, as will be discussed in each section. Please note that the axes on each individual profile for meal points consumed are on student-specific scales, therefore it is not useful to compare these graphs across students.

1.7.1 Complete smoothing

Figure 1.4 below shows just a few of the 46.4 percent (2075 individuals) whose behavior can be characterized statistically by complete smoothing. The estimated standard deviation on the error distribution is shown at the top of each chart, and is the only parameter estimated in this model, since complete smoothing makes a precise

prediction of the mean optimal path. Note that this parameter value is often fairly high, reflecting the fact that week to week consumption is volatile. However, recall the definition of smoothing which is that students only need to adjust their consumption on a weekly basis in the correct direction and on average by the correct amount.

In each student's graph, the (blue) line with open circles represents a student's actual consumption, the (red) dots represent the predicted complete smoothing path, and the (light blue) dashed line represents the consumption needed to follow the school's suggested plan. Every time the red dot fills the blue circle, the student has hit the predicted smoothing path. The shape of the school's plan as well as the complete smoothing path differs for each individual, because they depend on the previous errors made by each individual student. For students classified as perfect smoothers, the number of times they over-consume relative to their predicted path is about the same as the number of times they under-consume.

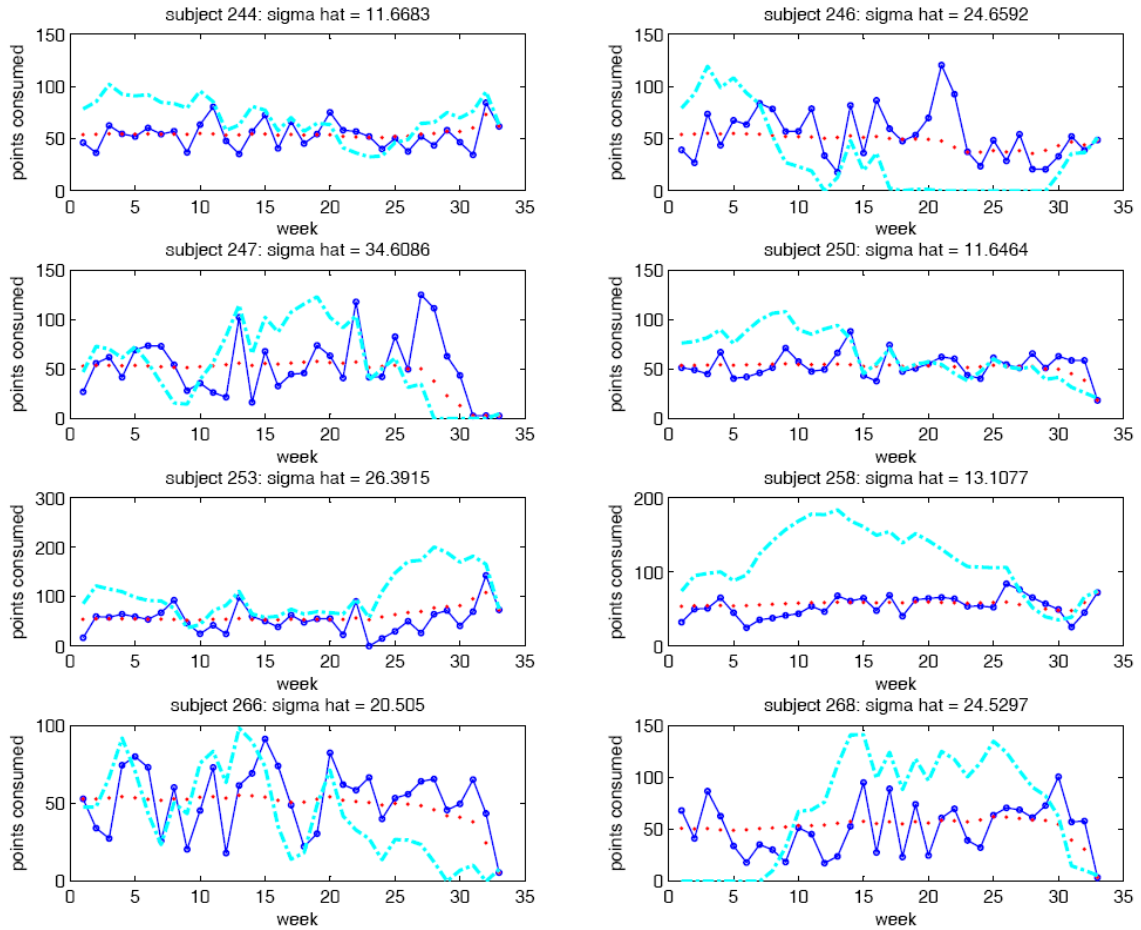


Figure 1.4: Examples of complete smoothing individuals

Subject 244 is a good example of a complete smoothing student. He or she managed to consume the complete smoothing target several times over the course of the year. Furthermore, when he or she fails to consume the smoothing target it is an over-consumption about as often as an under-consumption. He or she does not excessively deviate from the school's suggested plan either, although smoothing clearly provides the better fit, particularly in the beginning of the year. Another interesting example is Subject 258, who also hit the smoothing target or very close to it, several times during the year. Over the first several weeks however, he or she grows increasingly farther from the

school's suggestion, and the inflexibility of the school's plan clearly cannot compete with the smoothing prediction.

1.7.2 Biased smoothing

Figure 1.5 below shows some of the 20.6 percent (or 922) of participants in the sample who were best described as having a systematic bias to their consumption relative to complete smoothing, and furthermore I was able to reject that there was any point where the student stopped having that bias. In other words, throughout the year they had been adjusting their spending in the correct direction relative to the smoothing incentives, but also tended to miss their newly calculated weekly consumption target by a fixed fraction. The fraction of the complete smoothing plan consumed λ_n , is significantly different from 1 at the 95 percent level.

The red dots are the predicted consumption path under the estimated bias λ_n , while the green dashed line shows the predicted path for complete smoothing. In each case, we can see that the errors are much better centered around the red dots compared to the green dashed line. It is clear that students in this category are reacting to the smoothing incentive and expiration of meal points, but on average they do not manage to adjust fully.

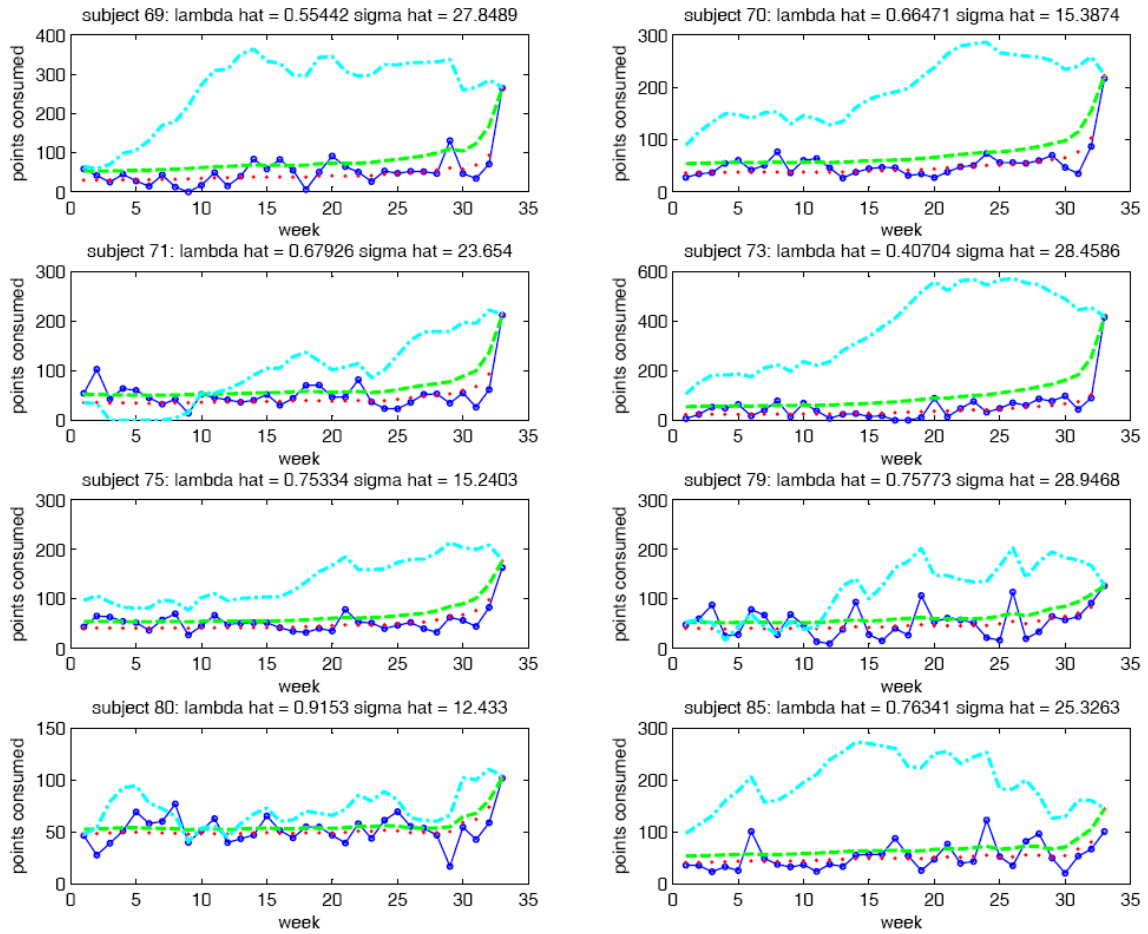


Figure 1.5: Examples of biased smoothing individuals

Subject 69 provides a typical example of a student consuming according to biased smoothing. He or she consumes approximately 55% of the complete smoothing solution throughout the year, with the red dots representing the predicted consumption level with this bias. The upward trend that is predicted by the complete smoothing path is in fact being tracked as time goes on, but only at a fractional level. In the final week, the budget balance is satisfied when he or she consumes the remaining meal points in his or her account (recall that satisfying the budget at the last opportunity is not being ‘counted’ as smoothing behavior). The other example students in this category show a very similar story.

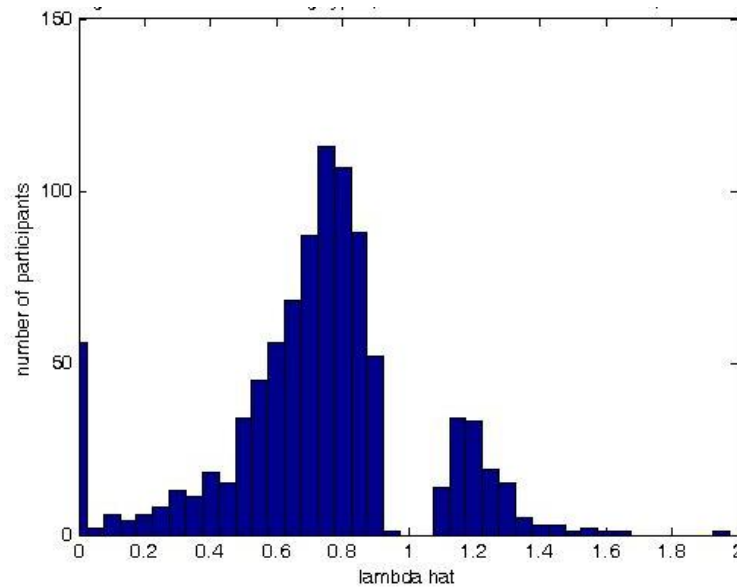


Figure 1.6: Biased smoothing types, distribution of fractions lambda, (N = 922)

Figure 1.6 shows the full distribution of estimated bias parameters, λ_n for students in this group. The distribution is skewed to the right, with about 86 percent of students in this category having a systematic under-consumption bias and only 14 percent having an over-consumption bias. The skewness in the distribution of λ_n seems to support the idea that the initial allotment of points was larger than naturally desired for many students.

A technical point concerns the mass of over 50 students with λ_n very close to zero. The problem here is a very high personal variance of errors around the model that prevents accurate estimation of any parameters. Any large deviation from the model relative to the individual's typical consumption will interfere with the estimation in this way. In particular for the 50 individuals that were grouped into this category, the most common source of the poor fit was due to the student foregoing hundreds of meal points instead of satisfying the budget at the end of the year as the model predicts, resulting in a large end-of-year error term.

1.7.3 Biased smoothing with learning

Figure 1.7 shows some of the 33 percent (or 1474) of students who were best classified as showing learning in their consumption choices. The particular week that complete smoothing is estimated to occur k_n , is shown in the title of each graph. Based on the classification criteria, these are students for whom I could not reject the hypothesis that learning occurred. Recall the rationale for this method of grouping was to give students the benefit of the doubt towards the most rational prediction of the model, even if this rationality occurred late in the year.²⁴

The red dots show the predicted path under this learning model, the blue line once again shows the student's actual consumption, and the green dashed line shows the complete smoothing alternative. Beginning with week k_n , the learning path coincides exactly with the complete smoothing path by definition. The school's suggested plan is omitted from most of these graphs since the prediction of that model tends to be especially far off from the prediction for this model, which interferes with the useful scaling of the graphs.

²⁴ I also considered a variation of this model in which students consumed at a flat level before week k_n , and found about 16% of all students fitting this alternative characterization. It is clear from these modeling experiments that a significant fraction of students are well-described by the hypothesis of having a better picture of the optimization problem later in the year than earlier in the year.

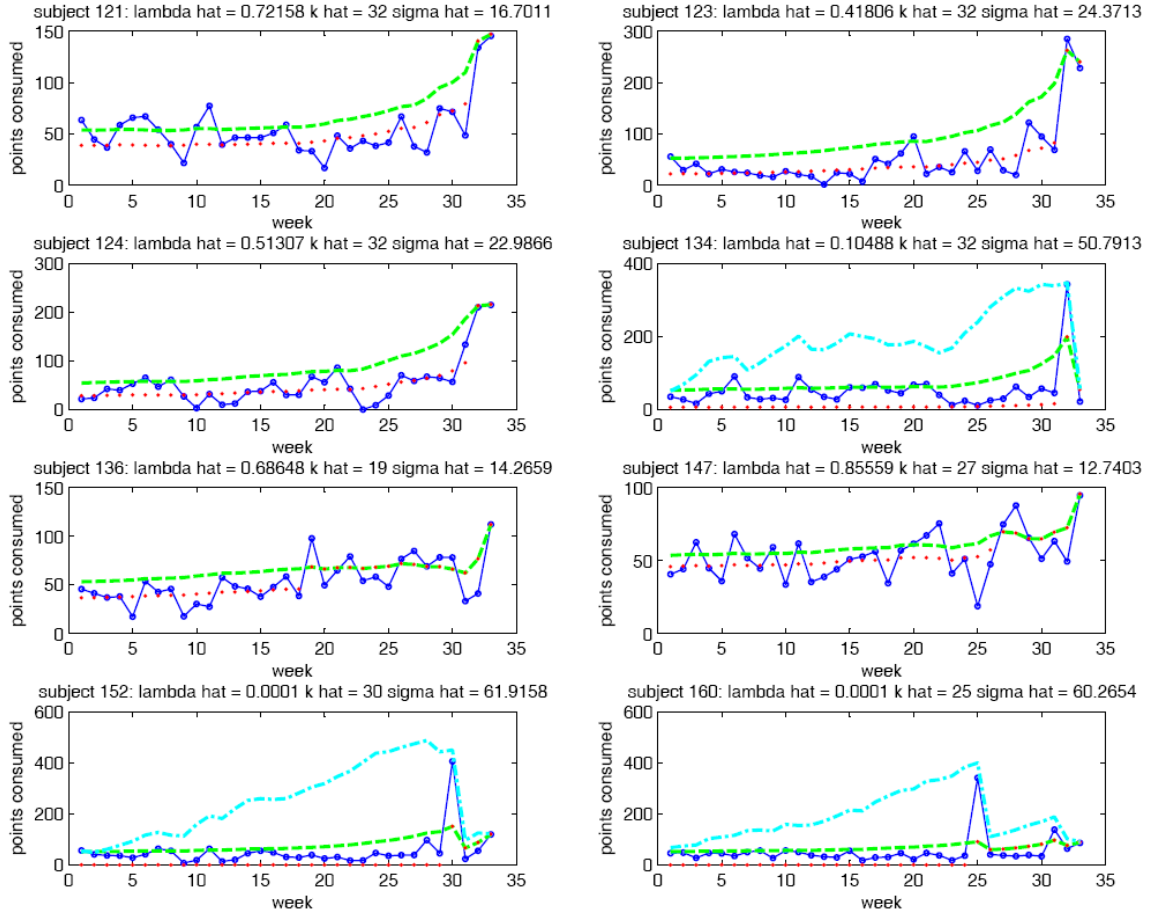


Figure 1.7: Examples of biased smoothing with learning individuals

Subject 121 provides a good illustration of the learning model. Up until week 32, the student consumed at approximately 72% of the complete smoothing solution, tracking the complete smoothing path, but better described as only a fraction of it. In week 32, he or she suddenly adopts the smoothing solution, dividing his or her remaining meal points nearly evenly over the final two weeks. Subjects 123 and 124 show similar stories.

Subjects 136 and 147 are examples of students placed into this category by virtue of not being rejected, since their learning was much more subtle. In both cases they appear to be following smoothing much better after k_n than beforehand, albeit not as drastically as some of the others.

Finally subjects 134, 152 and 160 are good representatives of another recurring phenomenon in the meal plan, students making a large pre-emptive splurge purchase before the end of the year, most commonly in the retail category. Students showing this type of behavior are not very well explained by any of the models under consideration because they do not consistently follow any single plan. Instead, they appear to be using the school's guidelines (shown for these specific students using the light blue dashed line) in a one-shot manner to figure out how many meal points they should have left by that particular week in the program, and making their retail purchases accordingly. Due to this one-time 'catching up' behavior, they best classify as learning even though they are not learning smoothing so much as learning to catch up to the recommended levels. As a result of this large deviation from any of the smoothing models, the bias parameter λ_n is unable to be estimated accurately and therefore comes out as the minimum possible value allowed. As can be seen from the distribution of estimated bias terms in Figure 1.8, these individuals comprise about 16% of all students in the learning category or 5% of the total sample.

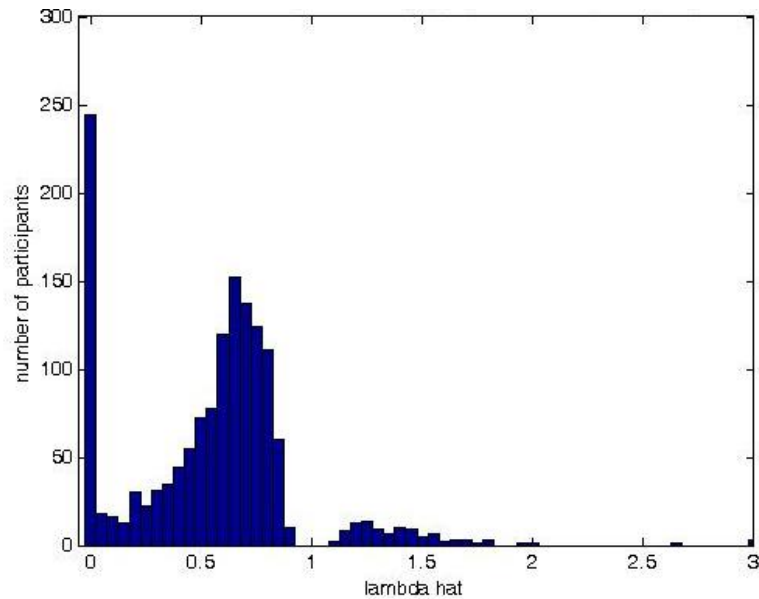


Figure 1.8: Biased smoothing with learning types, fraction of optimal plan, (N =1474)

Figure 1.9 shows the distribution of weeks k_n in which optimization began for these students. K_n rarely occurs in the first 10 weeks of the year, due to the statistical penalty of the additional parameters associated with learning compared to smoothing. The mode of the distribution is week 32, the next-to-final week of the meal plan (recall that learning in week 33 is not counted). The remaining individuals are distributed between weeks 10 and 31 in a slowly increasing frequency as the expiration date nears. It is clear that the expiration date serves as a strong incentive to adjust meal point use in the correct direction, and also that students tend to wait until expiration is fairly imminent before realizing this incentive.

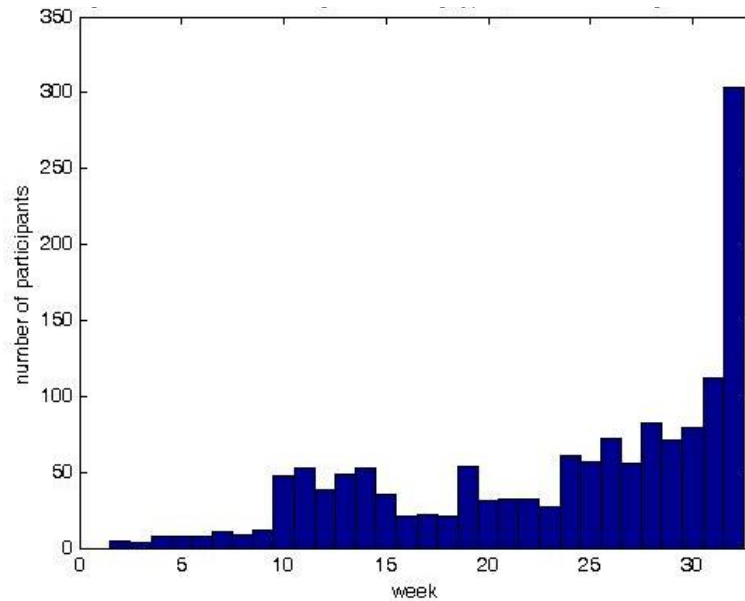


Figure 1.9: Biased smoothing with learning types, week of learning, (N = 1474)

1.7.4 School's suggestions

In the final sorting criteria, zero percent of students were best characterized as following the school's suggested calendar for meal point use. About 3 percent of students were unable to reject either the school's plan or its best competing model based on the Vuong test. However, visual inspection of these students' consumption profiles strongly suggests that this is due to the low power of the Vuong test, and not because the fits were anywhere close to one another. Following the school's plan imposes a strict consumption schedule for a student throughout the year, and only those who hardly ever made errors to the complete smoothing plan were remotely near the school's plan.

Although it proved unlikely to follow the school's plan for the entire year, a number of students appeared to use it as a benchmark late in the year when their consumption had gone astray, as evidenced by the fraction of learning students who had pre-emptive splurges in the previous section.

The rejection of the school's plan might be considered negative news if the concern is whether students were able to follow set guidelines, but it is actually a positive sign of students' economic intuition. This is due to the fact that any of the other models I considered require more economic rationality on the part of the student, compared to following the predetermined plan.

I also estimated a hybrid model of the school's suggestion, in which students waited until some intermediate week before adhering to the school's plan. Overall, this model fared about as poorly as the regular school's guideline model, because even if a student decided to catch up to the school's plan in a given week, going forward he or she tended to implement smoothing rather than continue following the school's plan. In other words, even students who had a late start in realizing the planning problem were more interested in smoothing than sticking with the guidelines.

1.8 Conclusions

The meal plan provided a field environment to examine the types of deviations from consumption smoothing that can occur when individuals are given the freedom to deal with their dynamic allocation problem the way the 'solution' comes to them naturally. Although students in the program did not have the freedom to choose the initial allotment of meal points provided, the data show that students typically only ate fewer than half of their meals using the meal plan, refuting the idea that 1800 meal points exceeded their upper bound for meal consumption. A key insight of the analysis is that these systematic smoothing deviations were prevalent even in an environment that was completely predictable and unchanging. Since students who either finished their meal

points early or purchased more points were excluded from the sample, the percentages of people fitting each of the models should not be taken as representative of any population, but only as a basic assessment of the major ways in which points were spent in this setting.

The choice students faced was how to allocate their stock of meal points over the school year. The intuitive prediction and one most commonly cited in the survey is week-by-week complete smoothing which is easy to implement, does not require any complex calculations, and is actually less rigid than following the school's suggested plan. The results from model estimations show that nearly half of students in the sample could be characterized as following this simple model of spending. Another thirty percent of participants were able to switch from whatever allocation they were initially implementing (typically under-spending), to one which was in line with spreading points evenly over remaining weeks. The fact that this many people experienced a discrete jump in spending near the end of the program is informative. While there might be nothing economically noteworthy about a student using as few points as he needed for the entire year, then using his remaining meal points to throw a party, the students who learned had (in estimation) a series of roughly equally sized "parties" over the course of two, three, or more weeks – this strongly suggests that at least some students "woke up" to the merits of smoothing their spending very late into the program.²⁵

²⁵An alternative explanation is that students fully anticipated having a shift in utility of meal plan items, so that marginal utilities of spending levels before and after the shift were the same in spite of the large increase in spending levels. With the current data, I cannot rule out this alternative explanation – however it seems unlikely that the concentration of utility shift points would be so heavily concentrated across students near the final week coincidentally. A more likely explanation is that the deadline (and posted

Since I censored the sample to exclude students who finished their meal points early, an opposite but similarly delayed reaction might have occurred for those who “woke up” and discovered that they had too few meal points remaining relative to the number of remaining weeks. However, since these students would have had to face the additional challenge of deciding whether and how many more meal points they would purchase, theirs is a more complicated decision that the data do not have enough information to address. Nevertheless, the results of this study should not be interpreted as though a similar but opposite effect did (or did not) exist for those who demanded more meal plan items than the initial endowment of points.

Two alternative models that I was unable to consider involve uncertainty and habit formation. The decision to keep models as simple as possible instead of exploring these possibilities further is mostly due to the limitation that each student only has 33 observations and the fact that the noise in the data even around the very simple models considered is quite high. Uncertainty is a critical factor in the outside world, but it is difficult to imagine that it is one of the overriding influences in the meal plan for the subset of participants examined, given the static environment.²⁶ Habit formation is another plausible hypothesis, however once again the data are too limited and noisy to pursue verifying or rejecting the existence of habit.

The analysis leaves several questions unanswered, in particular, what accounts for the fixed-fraction bias in smoothing? Consuming a fraction of the complete

reminders) brought on students realizations of wanting to smooth the remaining points. However, I cannot show using the data alone that this was the single driving explanation.

²⁶A number of the survey responses expressed that uncertainty about own future behavior was a concern, although the implications of this uncertainty would be different from uncertainty about income shocks, for example.

smoothing plan is a convenient general way to describe bias, but it does not specify whether it is due to a misconception about the number of meal points left, the number of weeks left, or some other factor outside the model. Further research using other data sources, or perhaps laboratory experiments related to time and wealth awareness, might be better equipped to clarify this issue.

Another unanswered question is, for students who switched from biased smoothing to complete smoothing, what factors prompted them to do so? The high concentration of learning weeks towards the end of the year suggest that the expiration date of the meal points played an important role; even if students discovered their intertemporal problem late, they eventually came around to the conclusion that smoothing their spending was appealing. This was clear from the drastic increases in retail sales over the last three weeks of the program. External factors that likely played a role include the reminder signs posted in dining facilities in the final weeks of the year, as well as peer influences or reminders. While this study did not attempt to find the source of learning, further work on what factors could most capably promote consumption objectives would be useful from a policy standpoint.

1.8.1 Policy implications for meal plans

The most immediate policy question is whether there is room for improvement in university meal plan design. Universities face a dual responsibility of offering flexibility in dining choices and providing healthy options for students, many of whom are new to making their own food choices. Therefore, it is understandable why the college would mandate that residents purchase a sum of meal points that would cover approximately

half of a student's meals over the year. In addition, in most campuses of large public universities, the residence hall program must compete with fast food franchises that are located in or near campus. This is presumably beneficial for students since the meal plan has to then make an effort to sell foods that are comparable in appeal.

The tally of meals suggests that students eat an average of about one meal a day inside the meal plan. Combined with the heavy concentration of retail purchases in the final weeks of the program, this suggests that both the university and students could benefit from quarterly plans instead of the year-long plan.

With quarterly expiration dates, the concentration of retail sales should at least disperse evenly between the three end-of-quarter deadlines, allowing students to use to retail items purchased at the end of two quarters during the next quarter. In a yearly plan, students may be hesitant to use their points earlier in the year for goods which they may feel are incorrectly priced, however they realize later on that this is their best option. Another possible effect is that the shorter time horizon may make students more aware of the meal allocation task, resulting in a lower sum of total retail sales. Finally, having three shorter dynamic tasks a year would most likely lead to learning as found by Brown et al (2009), and more ideal meal point allocations in the second and third quarters of the year. However, there may be negative returns to shortening the expiration horizon to less than one quarter – at that point it may actually become burdensome for students to spend their points on schedule, resulting in overall dissatisfaction with the meal program.

If meal points were in fact inefficiently allocated over the year, then opportunities may exist to correct this less than ideal allocation. The most obvious one is for students

with too many meal points remaining near the end of the year to buy meals for people who had run out of meal points but still needed to consume dining hall food. Anecdotal survey evidence suggests that buying for friends was one strategy that students used to spend extra meal points. However the retail sales data suggests that this was not the predominant way that most students smoothed their spending. A major barrier to swapping meal points is inconvenience, as well as finding matches between students willing to pay for additional meal points and those willing to sell them. Facilitating such transactions would most likely lead to reduced revenues for the university, and reduced success in the school's objective of having students eat healthily over the year.

Another practical possibility is for the school to make arrangements with on or off-campus "fast food" establishments to expand the pool of dining facilities where students can use their meal points. The benefit is that the university plan could then capture two or three of a student's meals on average rather than just one meal. By reducing the number of non-meal plan affiliated establishments, students would then have to eat almost all of their meals within the plan because other options would be geographically far away and significantly less convenient. Such an arrangement would be beneficial for the fast food establishments since all students living on-campus are required to subscribe to the meal plan, and some of the existing food demand could be shifted towards their products. It could be beneficial to students by offering more choice on how to spend their meal points. The university could also benefit by requiring that students purchase a greater number of meal points with the housing contract. Nutritional objectives however, might fall by the wayside as a result.

1.8.2 General policy insights

Do the results lead to any general policy insights beyond university dining? If the data are taken as a single example of how people tend to react to particular structural program rules, then insights from the dining program might be extrapolated to other settings.

1. *Relatively few meal points were actually left to expire, and deadlines matter for the timing of expenditures.*

A candidate related policy is consumer stimulus checks or tax rebates designed to promote spending. The meal plan data showed that although students often were frequently slow in using their meal points, they very seldom left any money on the table, spending large amounts in the final week if necessary. This would suggest that stimulus in the form of debit cards with an expiration date might be an effective way to hand money back to households while ensuring that it is spent rather than saved. The meal plan suggests that even if the debit cards are restricted to a confined basket of consumption (ex. food, entertainment, technology), consumers would find a way to use it – although if the amount is more than needed for a single purchase, or if consumers prefer to spend on a different category first, we might expect a similar last minute spending surge near the deadline.

2. *A plausible interpretation of a prominent spending pattern in the data is that students realized late into the program that they were interested in smoothing future spending.*

Similar spending/saving patterns may result when individuals realize late in life that they have saved too little or too much relative to their life expectancy. Although lifecycle consumption is a more complicated problem due to uncertain individual lifetimes, the fact that non-smooth spending occurs under the simplicity of the certain horizon meal plan suggests that fundamentally similar distortions also exist in lifecycle consumption, in addition to the standard complications such as lifetime/income uncertainty, changes in consumption demand depending on family size, etc. Results support the idea that the young do not plan very well for the future because much of the future is beyond their planning horizon (see Caliendo and Aadland, 2006).

3. Students rejected the school's guidelines on how to spend their meal points.

The results imply that “light suggestions” on how to make intertemporal allocations which are not enforced by deadlines or incentives, are rarely followed. This supports the findings of Choi, Laibson, Madrian and Metrick (2001) that financial seminars in the workplace had only a very limited impact on savings rates. This suggests that policies which only make recommendations or instructions for individual behavior without providing any material incentive or penalty for following them, are likely to have little or no impact.

Chapter 1 has been submitted for publication. The dissertation author was the primary investigator and author of this paper.

Appendix 1.A: Survey responses

I am doing a study about students' experiences using the meal plan. I would really appreciate if you could take a few minutes to think carefully about, and answer the following questions.

1. Year & quarter entered _____ My major is _____ Gender _____
2. This is my (circle one) 1st 2nd 3rd 4th academic year on the meal plan for on-campus students (begins with 2100 points this year). Leave the rest blank if you are not currently on this meal plan.
3. Please describe how you typically decide how many meal points to spend on each visit to a dining facility.

4. Which of the following best describes how you intended to use your meal points? (check one and explain)
 - _____ I intended to spread my use of meal points evenly throughout the year.
 - _____ I intended to use more of my meal points earlier in the year than later in the year.
 - _____ I intended to use more of my meal points later in the year than earlier in the year.
 - _____ I intended to use more of my meal points sometime in the middle of the year.
 - _____ I intended to use my meal points as necessary to buy whatever meal plan items I wanted, without considering whether I might have fewer or more meal points to spend in the future.
 - _____ None of the above (specify) _____.

Please explain why: _____

5. Which of the following best describes how you **actually** used your meal points? (check one and explain)
 - _____ I spread my use of meal points about evenly throughout the year.
 - _____ I used more of my meal points earlier in the year than later in the year.
 - _____ I used more of my meal points later in the year than earlier in the year.
 - _____ I used more of my meal points sometime in the middle of the year.
 - _____ I used my meal points as necessary to buy whatever meal plan items I wanted, without considering whether I might have had fewer or more meal points to spend in the future.
 - _____ None of the above (specify) _____.

Please explain why: _____

6. If at any point I realized I had more meal points left than I had desired, I responded by (check those that apply)
 - ☐ buying more expensive meals ☐ buying more pre-packaged goods ☐ eating more often in the dining halls
 - ☐ I never found myself in this situation ☐ Other (specify): _____
7. If at any point I realized I had fewer meal points left than I had desired, I responded by (check those that apply)
 - ☐ buying less expensive meals ☐ buying fewer pre-packaged goods ☐ eating less often in the dining halls
 - ☐ I never found myself in this situation ☐ Other (specify): _____
8. I was uncertain about whether I could actually afford to buy more meal points if I needed them. (circle one)
 - not at all a little somewhat quite a bit extremely this did not apply to me
9. I thought that my parent/guardian would use knowledge of my meal point usage to help decide how much spending money to give me. (circle one)
 - not at all a little somewhat quite a bit extremely this did not apply to me
10. The dining facility where I most frequently went to spend my meal points was (circle one)
 - A B C D E F G H I J Other (specify): _____

Thank you for participating! Feel free to write any additional comments on the back of the page.

I conducted a short survey of the students using the 2007-2008 meal plan during an introductory economics class, as well as an intermediate economics class.²⁷ The purpose of the survey was to gain insights as to how students thought about and approached their meal plan choices. The meal plan used by survey respondents was essentially identical to the one I empirically analyzed, the only difference being a higher initial allotment of points (2100 in 07-08 compared to 1800 in 05-06). The survey was conducted in the middle of week 32.

Students were told that I was doing a research project to understand how students spent their meal points over the course of the year. The survey was voluntary, but since class time was given for filling out the survey the participation rate was near 100 percent, excluding students who were not using and had never used the meal plan.²⁸ In the survey, students were asked to explain how they made purchase decisions during visits to the dining facilities and how they had intended to use their points over the year, using both multiple choice questions and free response. Almost all students answered the free response questions with some degree of detail. A copy of the survey is at the back of the paper.

Intended use of meal points

Of the sample of 81 students, 41 percent said that they intended to smooth their meal points evenly over the year, while about the same number (44 percent) said that they simply purchased ‘as necessary’ without considering the future. Of those who intended to smooth their meal points, the explanations given were consistent with the standard

²⁷ A spreadsheet of students' responses to the survey is available on request.

²⁸ A few students filled out the survey referring to previous years being on the meal plan.

economic rationale for consumption smoothing. Several participants gave descriptive reasons along the lines of, “So that I have enough to eat or buy throughout the year.”

An important insight regarding student intentions is that virtually no one expressed a desire to consume more initially and less later in the year, which is the typical prediction for standard long-run discounting.²⁹ If a time-consistent discount factor is less than one, then no matter when the participant is asked about his intentions, he should have had a desire to consume more meal points earlier on since utility earlier is always better than utility later. This suggests that the standard time-discounting framework is not the appropriate model for the meal plan setting.

The top rows of Tables 1s, 3s, and 4s show the break down of intentions for meal plan usage by gender (Table 1s), major (Table 3s), and number of years experience on the meal plan (Table 4s). Neither gender nor academic major exhibit differences across these demographic traits in terms of the desire to smooth consumption evenly compared to simply consuming ‘as necessary’. Table 4s however, shows a slight difference between the intentions of those who were on their first year using the meal plan compared to those who had already completed one year on the meal plan. A greater proportion (53 percent) of all first-time participants expressed an intention to smooth their use of points evenly compared to those who had already used the meal plan before (33 percent).

One possibility is that because of their experience in the previous year’s meal plan, repeat students did not have to place as much deliberate emphasis on smoothing but had instead developed an ‘automatic plan’ of what to do, which corresponded more closely

²⁹One person in the sample chose this as their intention, but did not elaborate in words.

with the ‘as necessary’ intention. For example, one participant who intended to eat ‘as necessary’ said, “From experience I learn(ed) that I will generally have leftover meal points.” Other rationales given for consuming on an as needed basis, included aversion to implementing budgeting when it came to a necessity such as food, and knowledge that they would always be able to buy more meal points if needed or buy food elsewhere.

Note that at least half of students who reported that they intended to spread point usage evenly, reported actually doing something else instead. Some of them gave circumstantial reasons for the discrepancy (for example, not having time to cook later in the year, or not having a means of going off campus early in the year). Others gave explanations that were less individual-specific. A couple of people cited the fact that they started buying more expensive items later in the year. A person who intended to smooth their point usage, but instead ended up consuming more in the early part of the year said, “It (dining hall food) was easily available and easier to overlook my preferred spending limit.”

Some intended smoothers ended up eating ‘as necessary’ instead. The explanations for this deviation ranged from “There are too many meal points,” to “There was a good variety of things I wanted.” One participant gave a more explicit reason, “I didn’t put too much thought into my ‘plan’ to spend all my meal points and I ended up with too many.”

Table 1s:		Which of the following best describes how you intended to use your meal points?						<i>N</i>
		spread evenly	more earlier in the year	more later in the year	more in the middle	as necessary	none of the above	
% of all males		40.0%	2.2%	8.9%	0.0%	44.4%	4.4%	45
		18	1	4	0	20	2	
% of all females		44.4%	0.0%	11.1%	0.0%	44.4%	0.0%	36
		16	0	4	0	16	0	
% of total		42.0%	1.2%	9.9%	0.0%	44.4%	2.5%	
		34	1	8	0	36	2	
		Which of the following best describes how you actually used your meal points?						<i>N</i>
		spread about evenly	more earlier in the year	more later in the year	more in the middle	as necessary	none of the above	
% of all males		22.2%	13.3%	13.3%	6.7%	37.8%	6.7%	45
		10	6	6	3	17	3	
% of all females		11.1%	8.3%	19.4%	5.6%	52.8%	2.8%	36
		4	3	7	2	19	1	
% of total		17.3%	11.1%	16.0%	6.2%	44.4%	4.9%	
		14	9	13	5	36	4	

N = 81

Table 3s:		Which of the following best describes how you intended to use your meal points?						<i>N</i>
		spread evenly	more earlier in the year	more later in the year	more in the middle	as necessary	none of the above	
% of all science majors		39.1%	0.0%	17.4%	0.0%	43.5%	0.0%	23
		9	0	4	0	10	0	
% of all economics majors		41.2%	0.0%	8.8%	0.0%	44.1%	5.9%	34
		14	0	3	0	15	2	
% of all other majors		45.8%	4.2%	4.2%	0.0%	45.8%	0.0%	24
		11	1	1	0	11	0	
		Which of the following best describes how you actually used your meal points?						<i>N</i>
		spread about evenly	more earlier in the year	more later in the year	more in the middle	as necessary	none of the above	
% of all science majors		13.0%	4.3%	21.7%	8.7%	47.8%	4.3%	23
		3	1	5	2	11	1	
% of all economics majors		17.6%	11.8%	14.7%	5.9%	41.2%	8.8%	34
		6	4	5	2	14	3	
% of all other majors		20.8%	16.7%	12.5%	4.2%	45.8%	0.0%	24
		5	4	3	1	11	0	

N = 81

Table 4s:		Which of the following best describes how you intended to use your meal points?						<i>N</i>
		spread evenly	more earlier in the year	more later in the year	more in the middle	as necessary	none of the above	
first year on meal plan		52.6%	0.0%	10.5%	0.0%	34.2%	2.6%	38
		20	0	4	0	13	1	
second (or more) year on meal plan or had meal plan in a previous year		33.3%	2.4%	9.5%	0.0%	52.4%	2.4%	42
		14	1	4	0	22	1	
		Which of the following best describes how you actually used your meal points?						<i>N</i>
		spread about evenly	more earlier in the year	more later in the year	more in the middle	as necessary	none of the above	
first year on meal plan		23.7%	5.3%	18.4%	7.9%	39.5%	5.3%	38
		9	2	7	3	15	2	
second (or more) year on meal plan or had meal plan in a previous year		11.9%	16.7%	14.3%	4.8%	47.6%	4.8%	42
		5	7	6	2	20	2	

N = 80

Visit-by-visit decisions

Another topic of interest in the survey are the responses to question 3, which asks students to describe in their own words how they typically decide how many meal points to use on each visit to a dining facility. Responses broadly fell into one of two categories, those who used a price-based rule-of-thumb approach and those who simply let their appetite dictate how many meal points to use. For some examples,

- Depends on what I feel like eating, but (I) try not to spend more than 6 meal points at each visit.
- I just buy what I want to eat.
- I don't really decide ahead of time how many meal points to spend.
- \$6 to \$7.
- (Depends on) how hungry I am and the type of food being served.
- I was told it covers 2 meals, so initially (I spent) \$9 for 2 meals, but food was expensive so I spend around \$7 to \$9 per meal based on nutritional or hunger needs.
- Meals needed divided by meal points is about \$5.50.
- I determine it by how many meal points I have left.
- Each time I go, I limit myself to about \$10.

The two categories did not always correspond to the intuitive answer on how they 'intended' to use their points. For example, several people who selected that they intended to consume 'as needed', yet on a visit-by-visit basis they cited specific dollar

amounts. In reverse, some of the people who intended to consume smoothly cited that they just purchased whatever they felt like on a visit-by-visit basis. This suggests that most students were at least semi-conscious of the incentives to smooth meal points while also not wanting to ration meal points too strictly.

Adjustment strategies

I was interested in how students made consumption adjustments when they realized that they had either more or fewer meal points remaining than desired, highlighting the endogeneity of the consumption decision. Table 7s shows percentages of students who used each of the listed strategies conditional on having found themselves in that situation at some point during the year. The realization of having too many or too few meal points left, is necessary to incite the desire to smooth.

First, it is worth noting that none of the students checked ‘I never found myself in this situation’ for both more points left than desired and fewer points left than desired, which suggests that the meal plan was salient enough for them to pay attention to how they were doing at least once in the year. 59 percent of survey participants found adjustment in both directions relevant, in other words, they at some times realized they had too many points left and at other times realized they had too few points left. 17 percent of participants found that never found themselves with too few points left, and 23 percent never found themselves with too many points left. Most students also selected more than one adjustment strategy, hence the numbers in Table 7s do not sum to anything in particular.

Table 7s suggests a few things about how students re-adjusted their consumption in response to being too low or too high in their meal point balance. First, buying pre-packaged goods is a common way to increase consumption levels (about as common as buying more expensive meals and eating more often in the dining halls), but cutting back on pre-packaged goods is not a very common way to reduce consumption levels. This may reflect the fact that many students consciously avoid purchasing non-prepared foods in the first place unless they expect not needing those points for prepared foods.

Table 7s: Adjustment strategies				
	<i>If at any point I realized I had more meal points left than I had desired, I responded by</i>			
	buying more expensive meals	buying more pre-packaged goods	eating more often in the dining halls	other
% of those who did not check "I never found myself in this situation"	40.3%	38.7%	35.5%	25.8%
N = 62	25	24	22	16
	<i>If at any point I realized I had fewer meal points left than I had desired, I responded by</i>			
	buying less expensive meals	buying fewer pre-packaged goods	eating less often in the dining halls	other
% of those who did not check "I never found myself in this situation"	46.3%	11.9%	53.7%	26.9%
N = 67	31	8	36	18

Table 8s: Re-adjustment realizations by consumption intention				
<i>intended to use meal points as necessary without considering the future</i>	had realizations of both too few and too many points left	never had realizations of too few points left	never had realizations of too many points left	never had any realizations
N = 36	50.0%	16.7%	33.3%	0.0%
	18	6	12	0
<i>intended to spread use of meal points evenly throughout the year</i>	had realizations of both too few and too many points left	never had realizations of too few points left	never had realizations of too many points left	never had any realizations
N = 34	73.5%	11.8%	14.7%	0.0%
	25	4	5	0

Another point to notice is that in addition to increasing or decreasing the frequency of visits to the dining hall, and choosing more/less expensive meals (which were used in about equal frequencies), many students suggested ‘other’ adjustment strategies. A common ‘other’ strategy for cases where a student had more meal points than desired, was buying food for friends or roommates who didn’t have enough. The most common ‘other’ strategies raised for cases where a student had fewer points than desired were, buying more meal points, eating elsewhere/cooking, or altering the composition of a meal to include less expensive items (such as free water).³⁰

Resource uncertainty and distorted incentives

Table 2s shows the distribution of total responses for questions related to overall resource uncertainty and distorted incentives for consuming optimally due to parental subsidies. 71 percent of participants said that they faced no uncertainty about being able to afford more points, or that the issue was not relevant for them. Another 20 percent of participants said that there was ‘a little’ or ‘somewhat’ of an uncertainty about the issue, and another 8 percent cited ‘quite a bit’ or an ‘extreme’ uncertainty.

Table 2s also shows the break down of students for different perceptions of parental monitoring. In total about 77 percent of respondents reported no perception that their parent used knowledge of their meal point usage to determine how much spending money to give them. Another 16 percent cited ‘a little’ or ‘somewhat’ of such a perception, and about 8 percent responded with ‘quite a bit’ or ‘extremely’.

³⁰ Some of the 'other' responses could technically be regrouped under one of the suggested categories.

Table 2s: I was uncertain about whether I could actually afford to buy more meal points if I needed them						
	not at all	a little	somewhat	quite a bit	extremely	this did not apply to me
% of total	40.7%	9.9%	11.1%	3.7%	3.7%	30.9%
	33	8	9	3	3	25
I thought that my parent/guardian would use knowledge of my meal point usage to help decide how much spending money to give me						
	not at all	a little	somewhat	quite a bit	extremely	this did not apply to me
% of total	45.7%	9.9%	6.2%	4.9%	2.5%	30.9%
	37	8	5	4	2	25
N = 81						

Table 5s (top) shows the distribution of consumption intentions for various degrees of reported budget uncertainty. There are not enough observations outside of ‘not at all’ or ‘this did not apply to me’ to infer much about the different categories. However the general pattern that people either intended to spread their point use evenly, or as necessary, still holds.

Table 5s: Uncertain resources							
<i>Which of the following best describes how you intended to use your meal points?</i>							
<i>I was uncertain about whether I could actually afford to buy more meal points if I needed them</i>	spread evenly	more earlier in the year	more later in the year	more in the middle	as necessary	none of the above	N
"not at all" or "this did not apply to me"	37.9% 22	1.7% 1	13.8% 8	0.0% 0	43.1% 25	3.4% 2	58
"a little" or "somewhat"	41.2% 7	0.0% 0	0.0% 0	0.0% 0	58.8% 10	0.0% 0	17
"quite a bit" or "extremely"	83.3% 5	0.0% 0	0.0% 0	0.0% 0	16.7% 1	0.0% 0	6
<i>Which of the following best describes how you actually used your meal points?</i>							
<i>I was uncertain about whether I could actually afford to buy more meal points if I needed them</i>	spread evenly	more earlier in the year	more later in the year	more in the middle	as necessary	none of the above	N
"not at all" or "this did not apply to me"	17.2% 10	8.6% 5	19.0% 11	5.2% 3	43.1% 25	6.9% 4	58
"a little" or "somewhat"	17.6% 3	5.9% 1	11.8% 2	5.9% 1	58.8% 10	0.0% 0	17
"quite a bit" or "extremely"	16.7% 1	50.0% 3	0.0% 0	16.7% 1	16.7% 1	0.0% 0	6
N = 81							

Table 9s shows a decomposition of budget uncertainty and parental monitoring reported intensities separated by whether the participant reported an intention to smooth use of points, or an intention to consume 'as necessary'. The table shows that there is very little difference in reported impact of these two external factors in their strategy for using up their points. (ie. it is not that those who consumed 'as needed' were much more likely to report parental monitoring, or that those who wanted to smooth were much more likely to report budget uncertainty) In other words, students appear to have an intention on how to use their meal points that is not dependent on these two factors. This could be due to the fact that a large number of points are pre-attached to the housing contract, so obtaining more points is not a great concern for most students.

Table 9s: Uncertain resources and parental monitoring, by intended use of points				
<i>Which of the following best describes how you intended to use your meal points?</i>				
<i>I was uncertain about whether I could actually afford to buy more meal points if I needed them</i>	spread evenly		as necessary	
not at all	12	35.3%	14	38.9%
a little	3	8.8%	5	13.9%
somewhat	4	11.8%	5	13.9%
quite a bit	2	5.9%	1	2.8%
extremely	3	8.8%	0	0.0%
this did not apply to me	10	29.4%	11	30.6%
	<i>N = 34</i>		<i>N = 36</i>	
<i>Which of the following best describes how you actually used your meal points?</i>				
<i>I thought that my parent/guardian would use knowledge of my meal point usage to help decide how much spending money to give me</i>	spread evenly		as necessary	
not at all	12	35.3%	17	47.2%
a little	5	14.7%	3	8.3%
somewhat	2	5.9%	2	5.6%
quite a bit	2	5.9%	2	5.6%
extremely	1	2.9%	1	2.8%
this did not apply to me	12	35.3%	11	30.6%
	<i>N = 34</i>		<i>N = 36</i>	

Students with budget uncertainty might be prime candidates for ‘more (point usage) later in the year’ due to precautionary savings motives. However, for students who wanted to consume more later in the year, some of whom did express a precautionary motive in their own words, budget uncertainty was not the primary motivation (all eight of them said uncertainty about their budget was irrelevant). Instead, the precautionary motive seems to be driven by an uncertainty about their own behavior combined with a desire not to purchase any additional points. As one student said, “I was not sure how I would do, so in order to last,...I wanted to either be on track or have saved meal points.”

Of those who responded that they believed to some degree that their parents used their meal plan usage to determine an additional resource allowance for them, we again see spreading point use evenly and consuming points as needed as the main intended

plans. In terms of actual use of meal points, the small number of people selecting ‘quite a bit’ or ‘extremely’ makes it difficult to infer anything, but the most common choice was to consume ‘more earlier in the year’ (includes those that intended to consume as necessary). It suggests that for this fraction of students, the knowledge that their parent used their meal point spending to gauge their resource needs, may have contributed to spending relatively more in the beginning. However, none of the students in this category mentioned their parents in the explanations, instead citing either specific circumstances or that they simply “ran out”.

References:

- Angeletos, G., Laibson, D., Repetto, A., Tobacman, J., Weinberg, S., “The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation” *Journal of Economic Perspectives*, Vol. 15, No. 3 (Summer 2001) p. 47-68.
- Brown, A.L., Chua, Z.E., Camerer, C.F., “Learning and Visceral Temptation in Dynamic Savings Experiments” *The Quarterly Journal of Economics*, February 2009, Vol. 124, No.1, p. 197-231.
- Caliendo F., Aadland, D. “Short-term planning and the life-cycle consumption puzzle” *Journal of Economic Dynamics and Control*, Vol. 31, Issue 4, April 2007, pp. 1392 – 1415.
- Camerer, C.F., Chua, Z.E., “Experiments on Intertemporal Consumption with Habit Formation and Social Learning” Working paper. August 2009.
- Choi, J.J., Laibson, D., Madrian, B.C., Metrick, A., “Defined Contribution Pensions: Plan Rules, Participant Decisions, and the Path of Least Resistance” *Tax Policy and the Economy*, 2002.
- Dellarocas, C., “The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms” *Management Science*, Vol. 49, No.10, pp.1407-1424.
- Goldman, E., “Testing efficient market hypothesis for the dollar-sterling gold standard exchange rate 1890-1906: MLE with double truncation” *Economics Letters*, 69 (2000) p. 253 – 259.
- Mastrobuoni, G., Weinberg, M., “Heterogeneity in Intra-Monthly Consumption Patterns, Self-Control, and Savings at Retirement” *American Economic Journal: Economic Policy*, forthcoming.
- Shapiro, J.M., “Is there a daily discount rate? Evidence from the food stamp nutrition cycle” *Journal of Public Economics* 89 (2005) p. 303-325.
- Stephens, M., “‘3rd of the Month’: Do Social Security Recipients Smooth Consumption Between Checks?” *American Economic Review*, Vol. 93, Issue 1, March 2003, pp. 406-422.
- Vuong, Q.H., “Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses” *Econometrica*, Vol. 57, No. 2 (March 1989), p. 307-333.

Chapter 2: Slot Machine Stopping Decisions: Evidence for Prospect Theory Preferences?

2.1 Introduction

Do slot machine gamblers behave consistently with the prospect theory value function? This paper presents new evidence on whether risky decisions can be most adequately accounted for with a neoclassical model of choice under uncertainty or a reference dependent model, using field data from a unique dataset of slot machine gamblers. While previous studies have tested for forms of reference dependence in the field, this study presents clear new evidence of irregularities relative to the standard neoclassical model from a previously untested setting, utilizing the features of slot machines that make them conducive to such tests.

I analyze data gathered over the course of one month at a casino using magnetic strip cards, which track nearly 2400 gamblers' activity for a single visit, recording each gambler's numbers of bets placed, wagers, and winnings. The data provide clear statistical evidence of reference dependence in winnings for slot machine gamblers that is consistent with the value function proposed by Prospect Theory (Kahneman and Tversky 1979; henceforth PT), rejecting implications of Expected Utility (henceforth EU) under the utility of gambling framework proposed by Conlisk (1993).

The workhorse neoclassical model EU, assumes that a person maximizes the mathematical expectation of his or her von Neumann-Morgenstern utility function defined over wealth levels. PT is the workhorse behavioral model, which assumes

that an individual's decisions under uncertainty are determined by preferences over distribution of total final winnings relative to a reference point such as the status quo ante or expected return – a feature known as *reference-dependence*. In addition, PT value depends on whether the level of winnings is a gain or loss relative to this reference point, with losses of a given magnitude lowering value by more than gains of the same magnitude raise it – the key feature of PT known as *loss aversion*.

In the analysis, I focus only on the shape of the value function in PT as compared to EU, ignoring the component of PT which allows nonlinear probability weighting (for example, overweighting small probability events). By using a simple form of EU as the representative of neoclassical choice under uncertainty and establishing it as the null hypothesis, I am unable to concretely rule out or accept the full range of non-expected utility neoclassical alternatives. Therefore, the empirical analysis of this paper might be best interpreted as testing for reference dependence (or lack thereof) in gamblers' valuation of winnings - a key component of PT which defines the shape of the value function, and fundamentally distinguishes PT from neoclassical models.

I assume that a gambler's decision to stop making bets is determined by two main factors: realized winnings as the PT or EU theories suggest, and an intrinsic utility of gambling derived from the anticipation of the bet's outcome as proposed by Conlisk (1993). The utility of gambling component is likely to decrease marginally in the long run as a result of gambling fatigue, but is unobservable in the data. My analysis assumes that the utility derived purely from pulling the slot machine lever is completely

independent of the outcome of each pull and is heterogeneous across gamblers.³¹ So each gambler has their own pace at which playing the slot machines becomes tiresome, but all gamblers are assumed to have approximately similar preferences over the wealth resulting from the gambles – in the case of EU, a slightly positive marginal utility of winnings so that winning is preferred to losing, but risk aversion over small stakes is minimal – and in the case of PT, a similar preference for winning over losing except that risk aversion is markedly higher near a particular outcome which gamblers consider the reference point. While the data do not have the capability of ruling out all variations of neoclassical models, they distinctly show a higher propensity to quit gambling at a reference point, and under a utility of gambling assumption of the form suggested by Conlisk (1993), support the break-even effect of Thaler and Johnson (1990) being greater than the house money effect.

The slot machine environment is a simple and relatively controlled setting to test for EU and PT in the field. The environment is constant since machines pay out with the same probabilities regardless of day of the week or time of day. Furthermore, the outcomes of individual bets are random and independent of previous bets, and there is no skill required.³²

Due to data limitations which do not allow observation of every single bet made, I consider only a gambler's total number of bets over all machines played and his or her total net winnings for the visit. Following the main analysis of Post, van den Assem, Baltussen, and Thaler (2008), I ignore the option value of continuing to make bets,

³¹ This rules out wealth-dependent utility of gambling, such as valuing the intrinsic act of gambling more while winning or losing.

³² See Zangeneh, Blaszczynski, and Turner (2007), Chapter 2.

assuming that gamblers make the decision whether to quit or continue myopically, one bet at a time, given that they have already decided to visit the casino (due to enjoyment of gambling, and other potential factors which may have led him there).³³

Making use of the fact that individual bet outcomes are serially independent, the central limit theorem predicts that regardless of the original distribution of returns for a single bet, for a sufficiently large number of bets taken, an individual gambler's sum of winnings up until that point should be distributed normally.³⁴ Therefore in the cross-section of individuals, if we assume that individuals quit approximately randomly (and independently of one another) as suggested by EU, then the empirical distribution of winning levels observed in the data for gamblers who took the same number and type of bets should also be normally distributed.

By extending this basic reasoning to allow for aggregation over bets and wager sizes, and supplementing with simulations for comparison purposes, the analysis shows that the distribution of actual realized winnings has substantially higher kurtosis (4th moment) than a normal distribution, or simulations accounting for gambler heterogeneity. Since excess kurtosis means a higher peaked probability distribution, this is an indication of the presence of the key feature of PT, loss aversion with respect to the reference point where individuals quitting decisions are disproportionately concentrated. This result confirms the predictions of PT in a dynamic setting, replicating Thaler and Johnson's (1990) break-even effect which predicts lower likelihood of quitting while losing relative

³³ See Barberis (2009) for a theoretical analysis of both the decision to visit the casino, as well as the gambler's behavior once inside the casino using a non-naïve PT framework.

³⁴ Appendix D discusses the speed of convergence of cumulative winnings to a normal distribution under reasonable assumptions on the underlying payout distribution. Although full convergence is not reached, a distribution that is not excessively different from normal is obtained after fairly few total bets taken.

to the reference point, and the house money effect which predicts the similar result for winning relative to the reference point.

In addition, the empirical distribution of winnings are more positively skewed than simulations suggest, meaning that gamblers were more likely to quit while being “ahead” (relative to the reference point) instead of quitting while “behind”. This is consistent with the hypothesis that the PT value function has the additional feature of diminishing sensitivity in the gains and loss segments, making winning individuals risk-averse and losing individuals risk-loving. So, while on the margin individuals are still most likely to quit near the reference point, they are even less likely to quit while behind compared to while ahead.

The rest of the paper is organized as follows. Section 2.2 provides a brief review of related literature. Section 2.3 describes the key features of EU and PT models of decision making under uncertainty and derives their implications for gambling behavior. Readers already familiar with EU, PT, and related empirical literature might skip to Section 2.3.3. Section 2.4 describes the data. Section 2.5 discusses the assumptions and empirical strategy. Section 2.6 reports the results. Section 2.7 concludes.

2.2 Literature review

Camerer (1998) surveys a number of field studies which have used PT to explain puzzles that are hard to reconcile in an EU framework. For example, Benartzi and Thaler (1997) proposed a PT explanation of the equity premium puzzle – the phenomenon of risk premiums on equity returns being systematically much higher than a reasonable risk-averse EU investor would require in order to hold equity rather than bonds. They

suggested that the premium is so high because it must compensate not only for investors' risk aversion, but also for their loss aversion. Camerer, Babcock, Loewenstein and Thaler (1997) (see also Koszegi and Rabin (2006)) considered the role of reference dependence and loss aversion in cabdrivers' labor supply decisions, finding that drivers worked fewer hours when average daily wages were high, and more hours when daily wages were low – a behavior contrary to predictions of classical intertemporal labor supply substitution, but consistent with loss averse preferences in daily income.

In additional support for loss aversion, Odean (1998) shows that investors, via a “disposition effect”, tend to sell stocks whose purchase prices allow them to realize a gain, significantly more often than otherwise equivalent stocks for which selling would yield a loss. He suggests a PT explanation for this systematically different treatment of gains and losses (also see Barberis and Xiong (2006)). Similar to Odean (1998)'s work on stock market investors, Genesove and Mayer (2001) show that condominium owners seeking to sell their properties but facing an expected loss in the market, posted asking prices that were systematically higher than owners of similar condominiums whose expected selling price would yield a gain.

An early, key experimental study by Thaler and Johnson (1990) extends PT by focusing on the impact of previous gains and losses on current decision-making under uncertainty, thus applying PT to an explicitly dynamic setting. Thaler and Johnson find that risk aversion as revealed by choices, tends to decrease after a previous gain, which they call the “house money effect”. They also find that risk aversion decreases after a

previous loss, provided that a gamble offers the opportunity for a person to break even, which they call the “break-even effect”.

Perhaps the most direct field evidence in support of Thaler and Johnson’s dynamic generalization of PT is Post, van den Assem, Baltussen, and Thaler (2008), who examine decisions of contestants in the television show “Deal or No Deal?” In the show, contestants go through several rounds of deciding between risk-free ‘bank offers’ and the pursuit of a final prize of unknown value in a suitcase. Post et al find that as information about the final prize is uncovered in each round of the game, contestants were more likely to reject generous risk-free offers (often exceeding the expected value of the final prize), after uncovering information that suggested either a very favorable or unfavorable final prize. In other words, contestants were less risk-averse when evidence supported the idea that their final prize was either very bad (“break even effect”) or very good (“house money effect”) relative to reasonable expectations – a pattern most easily reconciled using PT.

These studies make PT a plausible candidate to explain dynamic choice under uncertainty. However the issue remains controversial. For example, Farber (2005, 2008) questions Camerer, et al’s (1997) interpretation of cabdrivers’ labor supply as reflecting loss aversion, arguing that driver and day fixed effects account for most of the PT-like stopping behavior, and further that the reference point is difficult to pin down (see however Crawford and Meng (2008), who show that Farber’s data do in fact support a PT interpretation as proposed by Koszegi and Rabin (2006)). Barberis and Xiong (2006) raise questions about whether the “disposition effect” found by Odean (1998) necessarily

follows directly from PT. Finally, List (2004) compares EU and PT implications in a sports card trading market, finding that while most inexperienced consumers' behavior was consistent with PT, most experienced consumers behaved in accordance with EU.

2.3 Models of decisions under uncertainty

2.3.1 Expected utility

In the EU framework, a gambler enters the casino with a given level of wealth, which along with his utility function, determine how he views the gambles he is about to take. In most of the economics literature, this given level of wealth is interpreted as either lifetime or annual wealth. An EU individual gambles due to a combination of his utility of wealth preferences as well as his preferences over gambling, both of which interact to determine at what point he decides to stop. He stops gambling when his utility of gambling no longer compensates him adequately for the change in expected utility of wealth anticipated from the next bet.

EU normally assumes that utility of wealth is concave, often further assuming that the individual's level of absolute risk aversion decreases as wealth increases - so that increases in wealth, other things equal, make a person less averse to mean-preserving risks involving absolute changes in wealth. A typical neoclassical utility function looks like the one in Figure 2.1 below, which shows a "flexible expo utility function" (with parameters $\alpha = 0.5$ and $\beta = 0.5$). This family of functions allows the individual's risk-aversion to be flexible in both relative and absolute terms with respect to wealth.³⁵

³⁵ Flexible Expo utility nests Constant Relative Risk Aversion (CRRA) utility as the special case where $\alpha \rightarrow \infty$ and Constant Absolute Risk Aversion (CARA) utility as the special case where $\beta = 0$. The original

Attitudes toward risk in the EU framework are completely determined by a gambler's wealth level. For small stakes gambles such as slot machines, where the typical wager is only a fraction of a dollar at a time, a small amount relative to lifetime or annual income. Under these conditions, a differentiable von Neumann-Morgenstern utility function implies that people are approximately risk neutral. Further, even for variations in wealth as large as a *whole* visit's winnings or losses at a casino, which are on the order of less than \$1000 for 97% of people in the sample, significant variation in an individual's absolute risk aversion is unlikely. An EU individual with time-separable preferences will value an additional bet about the same everywhere within the range of wealth variation in a day at the casino, and therefore has no preference for quitting at one level of casino winnings over another. The EU individual's quitting decision can therefore be approximately interpreted as random with respect to winnings.

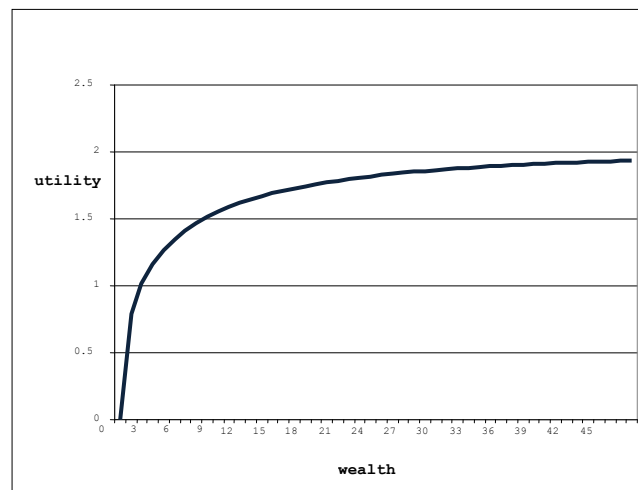


Figure 2.1: Example neoclassical utility function

formulation is credited to Saha (1993), and is utilized in Post et al. (2008) with the formulation $u(x) = [1 - \exp(-\alpha(W+x)^{1-\beta})]/\alpha$, also used in the above figure, where W is ex-ante wealth, and x is the net outcome of the gamble.

In this essay I avoid the scaling concern discussed by Rabin (2000) and Rabin and Thaler (2001) by confining the analysis to the small to moderate risks casino gamblers encountered in the dataset.³⁶

2.3.2 Prospect theory

Unlike EU, PT assumes that individuals evaluate the decision on whether or not to take a particular gamble based on their current position relative to some reference point. To close the model it is necessary to specify what determines the reference point. This has been done in various ways. For example the reference point might be determined by a person's expectations (Koszegi and Rabin 2007), or the status quo ante (Thaler and Johnson 1990). Since my econometric tests rely only on rejecting the normal distribution and not the exact value of the reference point, I do not attempt to estimate it, although its value can be inferred by examining where in the winnings distribution quitting tends to concentrate.

In our discussion of PT, an individual's valuation of winnings is referred to as the "value function", to differentiate it from EU's von Neumann-Morgenstern utility function. Recall that gains exceeding the reference wealth add value just as in the neoclassical

³⁶ Rabin (2000) and Rabin and Thaler (2001) take this argument a step further, arguing that EU cannot simultaneously explain people's attitudes toward small and large risks, because the moderate risk aversion commonly observed over small stakes gambles implies, in an EU framework, unreasonably high risk aversion over gambles of larger stakes. This is due to the fact that marginal utility would have to be dropping very steeply from the person's current level of wealth in the event of losing the small stakes gamble in order for them to reject it, which suggests that they will have to reject any gamble which poses a potentially larger loss. The specific example that Rabin and Thaler give is "Johnny is a risk averse, EU maximizer, who always turns down a 50-50 gamble of losing \$10 or gaining \$11. ... what is the biggest Y such that we know Johnny will turn down a 50-50 lose \$100, gain \$Y bet?" The answer is "Johnny will reject the bet no matter what Y is." The intuitive interpretation is that a person who finds the prospect of losing \$10 that unattractive even though he has the chance to win \$11 with the same probability, basically has to be so risk averse that he will turn down anything in order to guarantee that he doesn't lose \$100.

version of utility of wealth. However, losses below the reference level of wealth reduce value more than gains of the same magnitude increase it, producing loss aversion.

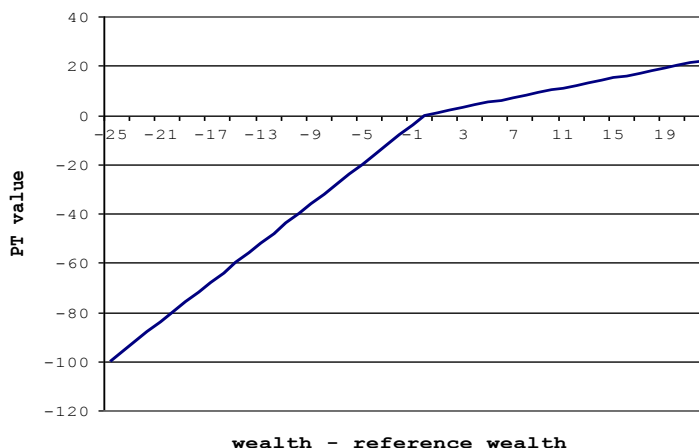


Figure 2.2: Example Prospect Theory value function

Figure 2.2 above shows the simplest possible example of a PT value function, one which assumes local risk neutrality in gains and losses away from the break-even point. Imagine a person starting at the reference point pictured above, who is offered a gamble with an outcome distribution that will take him above the reference point with some probability and otherwise force him below the reference point. Even though gains and loss segments are in isolation perfectly flat (risk neutral), near the reference point his value function is in fact very concave due to loss aversion. A person with this kind of value function over money is ‘first-order’ risk averse at the reference point, and approximately equally risk averse to small gambles and large gambles, thus avoiding

EU's weakness of not being able to reconcile real life decisions about risk on small and large scales.³⁷

The coefficient of loss aversion is defined as the ratio of the loss segment's slope to the gains segment slope near the reference point. In the example pictured, the coefficient is 4 (chosen to emphasize kink at reference point), meaning that losses are four times worse than gains are beneficial. Experiments typically produce estimates of the coefficient of loss aversion near 2, meaning that losses are twice as bad as gains are beneficial.

Kahneman and Tversky (1979) suggested that in addition to the loss aversion exhibited by the value function in Figure 2.3, the gains and loss segments may exhibit "diminishing sensitivity", in that deviations from the reference point have marginally smaller effects on value, the larger they are. This means the value function would be concave in gains and convex in losses as illustrated in Figure 3, making a person with these suggested PT preferences additionally ('second-order', as in EU) risk-averse over gains and risk-loving over losses.^{38 39}

³⁷ In first-order risk aversion the risk premium required to accept a fair gamble is proportional to the standard deviation of the gamble, whereas in second-order risk aversion the premium is proportional to the gamble's variance. See Segal and Spivak (1990) for a detailed discussion.

³⁸ A frequently used specification for the PT value function is $V(x|r) = -\lambda(r-x)^\alpha$ for $x \leq r$
 $(x-r)^\alpha$ for $x > r$ where r is the reference point, x is the ex-post level of wealth, and λ is the coefficient of loss aversion. α determines the curvature of the gain and loss segments. In Figure 2: $\lambda = 4$ and $\alpha = 1$, and in Figure 3: $\lambda = 4$ and $\alpha = 0.7$.

³⁹ Kahneman and Tversky (1979) also suggested a feature of prospect theory which we ignore in this paper: nonlinear probability weighting. Typically in PT, agents are allowed to weigh outcomes of gambles in ways that are not consistent with the actual odds, and furthermore the weights do not need to sum to one like actual probabilities. A common example of this weighting distortion is the overweighting of the tails of outcome distributions. Here I assume that people are fully rational about probabilities, so that the burden of explanation falls only on loss aversion and to a lesser extent, diminishing sensitivity.

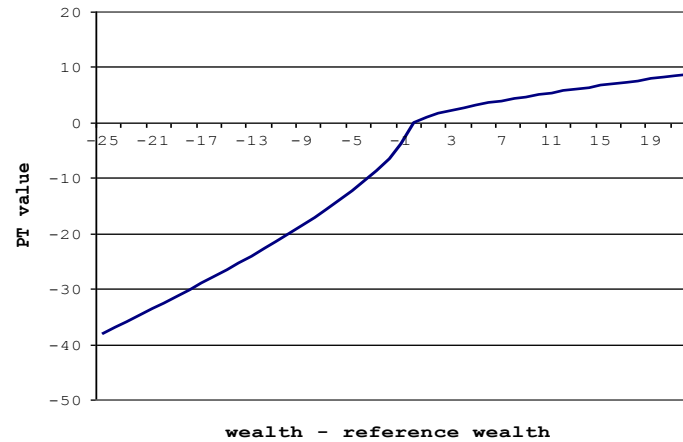


Figure 2.3 Example Prospect Theory value function with diminishing sensitivity

PT suggests that a gambler's risk aversion will differ systematically across the space of possible outcomes.⁴⁰ It follows that once the utility of gambling and heterogeneity among people are adequately controlled for as explained below, PT implies different likelihoods of a person quitting, depending which of the three regions they are currently in. Similar to the analysis of Post et al. (2008), I make a simplifying assumption that gamblers decide when to quit myopically, based only on the prospect offered by the next bet. However, allowing for more sophisticated gamblers who look several bets ahead would not change the qualitative results.

2.3.3 Predictions of PT and EU for slot machine gambling

Recall that the prediction of EU for the small stakes environment of slot machines is that gamblers quit approximately randomly, and in particular, there is no common reference level of winnings influencing the decision of when to quit. By contrast, PT

⁴⁰ EU technically does not rule out such variations, but does not particularly suggest them either in the sense that they would require unexplained and implausible variations in the coefficient of absolute risk aversion with wealth levels.

offers distinct predictions about relative likelihoods of quitting, depending how much the gambler has won or lost relative to the reference point.

Near the reference point a gambler will be first-order risk averse due to the concave kink in the value function, which unlike a differentiable vNM utility function, makes the gambler approximately equally averse to small and large risks. Therefore at net winnings levels near the reference point, a person is highly unwilling to take additional gambles, or equivalently more likely to quit.

A gambler who is far enough into the gains segment that he has little or no chance of crossing the reference point in the next bet, experiences Thaler and Johnson's (1990) *house money effect*, since he is risk neutral in the absence of diminishing sensitivity (Figure 2.2) or mildly second-order risk averse, (as in a standard EU analysis) with diminishing sensitivity (Figure 2.3). In either case, the house money effect becomes increasingly relevant the further the gambler is into the gains segment; but it may be enhanced for very large winnings by wealth-induced changes in absolute risk aversion. In other words, a gambler whose net winnings are far enough into the gains segment that he has little or no chance of crossing the reference point in the next bet (technically fully under the gambler's control by their choice of wager) will be more willing to take additional gambles, and therefore less likely to quit, than a gambler whose net winnings are closer to the reference point.

Following the same logic, a gambler who is far enough into the loss segment that he has little chance of crossing the reference point in the next bet, experiences Thaler and Johnson's (1990) *break even effect*, since he is risk neutral and willing to take risky bets

he would not take at the reference point, even in the absence of diminishing sensitivity (Figure 2.3). The presence of diminishing sensitivity (Figure 4) increases this effect further because it makes a person locally risk-loving in the region of losses. The overall result is that a person is least likely to quit when he is ‘losing’.

In practice, the slot machine environment has structural features that should provide a particularly conducive environment for observing Thaler and Johnson’s (1990) house money and break even effects, if these effects do exist. The first such feature is that a person has the ability to adjust the maximum amount he loses directly through his choice of wager. This means that if a person who is winning is truly loss averse, he is able to choose a wager size which will ensure that he will not cross his reference point. Secondly, from the perspective of gamblers who are losing, the notion of a “jackpot” provides the conditions that would allow a break-even effect to exist - a person who is currently losing always faces at least a very small probability of being able to break-even from the losses observed in this dataset in a single bet.

To summarize, the consequence of PT for stopping decisions is that controlling for the utility of gambling, individuals should be *most likely* to quit when they are near the reference point, *less likely* to quit when significantly above the reference point, and *least likely* to quit when they are significantly below the reference point. This is in stark contrast to EU theory, which assumes that preferences do not depend on a reference point, so that the distinction between gains and losses is irrelevant, further suggesting that risk attitudes should change very little over a series of small stakes bets.

2.4 Data

The data for the analysis are from a casino marketing program which tracks customer activity and offers participants reward points which can be redeemed for buffet meals, merchandise and other benefits.⁴¹ The structure and incentives of the program are similar to that of an airline mileage rewards program. Membership is free to anyone of the required age who provides a photo ID such as a driver's license, and their address.

After joining, customers are given a magnetic stripe card, which they insert into a slot machine as they play so that the casino can keep track of how much money they have wagered. Money put into the machines and tracked using their membership card are then converted into rewards points. The card itself does not actually hold a cash balance, and gamblers can technically play the machines without a card, however the rewards provide gamblers with an incentive to have the card track their activity. The card must be inserted into the machine while bets are being placed for gamblers to earn credits, helping to rule out multiple individuals playing simultaneously on the same card, or playing more than one machine at a time.⁴²

The dataset consists of one month's worth of new casino members each, for a single visit's worth of activity on the day that they sign up for the card. The data are aggregated at the machine level, rather than keeping track of every single bet. For each individual in the sample, the data covers slot machine activity for the entire length of visit commencing with membership enrollment - continuous gambling into the next calendar day is included. While there may be some differences in gambler characteristics

⁴¹ Since the data is proprietary, by agreement with the casino it cannot currently be made available publicly or to other researchers, nor can the casino be identified.

⁴² It is possible that a customer could have asked for a duplicate or replacement membership card, allowing for a spouse or friend to accumulate points on the same account. However, since these are first-time customers at this casino, the chance of card manipulation occurring on a large scale is unlikely.

depending the day of the week, or time of day, I do not focus on this and treat every observed customer identically.

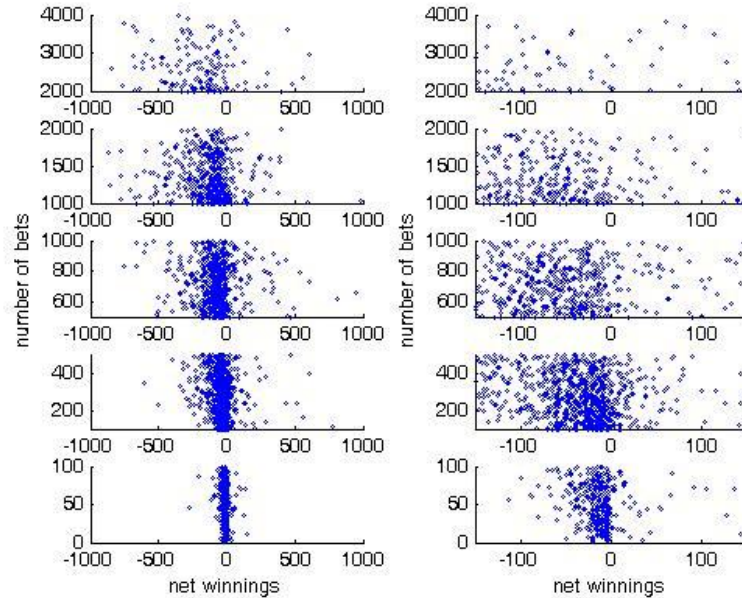


Figure 2.4a

Figure 2.4b

Figure 2.4a shows the distribution of the total number of bets (b) and total winnings (π) at quitting in (b, π) -space, with each gambler's experience represented by a single point. The x and y-axes are rescaled in a way that allows the points to be seen clearly (otherwise the points are too concentrated in a single area).⁴³ As is evident from the figure, there is considerable bunching at or just below zero net winnings, particularly for low numbers of bets, and less so for higher numbers of bets due to the negative drift to winnings as a result of the negative expected value. Figure 2.4b is identical to Figure 2.4a, except focused on a narrower range of net winnings. The figures show two conflicting tendencies in the data – a high likelihood of stopping near the break-even

⁴³ Seventy-nine out of the 2393 observations (a little over 3%) are outliers and omitted from Figure 2.4a for scaling purposes.

point (seen most predominantly for small numbers of bets), and the cumulative effect of negative expected returns on each bet on this tendency to stop at the break-even point (seen predominantly in high numbers of bets).

A PT framework, by allowing gamblers to react systematically differently to gains and losses defined relative to a reference point, might explain the concentration of stopping near zero total winnings in the data. By contrast, an EU model rules out any influence of a reference point by assumption, so the observed cluster could only be explained in EU terms by increases in individuals' coefficients of absolute risk aversion (making additional bets less appealing) that just happen to be especially concentrated near the break-even point despite individual heterogeneity in wealth – an unlikely explanation. This suggests that PT might explain the patterns in the data more gracefully than EU.

To make this point clearer, imagine that Figures 2.4 consisted of an extremely concentrated set of points forming a solid vertical line at the zero net winnings mark (a very extreme version of what the actual data hint at). This would mean that everybody in the sample had stopped at zero net winnings, no matter how many bets they had taken. For an expected utility framework to explain such behavior, every person in the sample would have to have a drastic upward spike in risk aversion at exactly zero net winnings, regardless of their wealth differences. Such uniform stopping would be a huge coincidence.

Figure 2.5 shows the data collapsed into histogram form. Clearly, the data show a high level of concentration near zero net winnings – this in itself could be interpreted as a

higher likelihood of quitting at a break-even point. However, in order to address this question rigorously, a more detailed analysis is needed, accounting for heterogeneity in utility of gambling as well as wager choices.

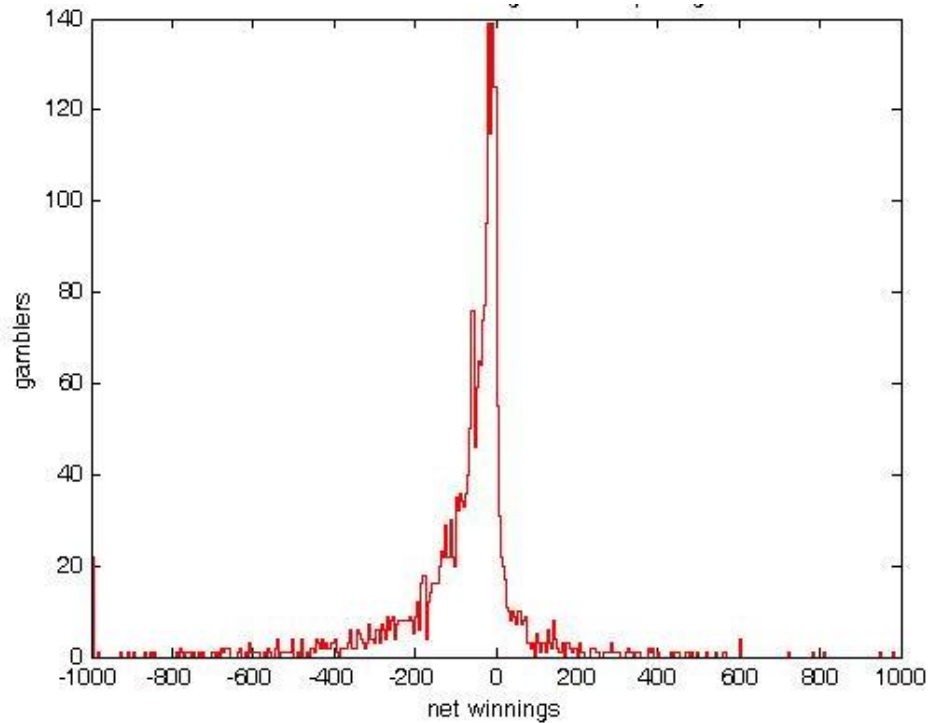


Figure 2.5: Distribution of winning levels at quitting

2.5 Predictions for individual stopping decisions

If bet-by-bet outcomes for each individual gambler were observed in the data, then a hazard model (see Kiefer 1988) could be used to test whether the probability of stopping was greater conditional on a reference level of winnings compared to other winnings levels. The hazard specification would estimate the probability that a gambler quits after a particular bet b , conditional on his cumulative winnings at that time π_b , given the fact that he had not quit in bets 1 through $b-1$ with the sequence of observed cumulative winnings after those particular bets. The resulting coefficient on winnings

would give the effect of winnings on the likelihood of quitting, which is the relationship of interest in distinguishing between EU and PT.

However, since the data do not allow observation of individual bet outcomes, the correctly specified hazard model cannot be implemented. Instead, I rely on a cross sectional approach which relates observed final winnings to the empirical likelihood of quitting, using completely random quitting as the null hypothesis representing EU. Statistically distinguishing between EU and the deliberate stopping near the reference point predicted by PT, relies on the independent nature of individual bet outcomes. Slot machines in the United States have payouts determined by random number generators resulting in bet-by-bet outcomes which are fully independent of any previous or future bet outcomes (Zangeneh, Blaszczyński, Turner, 2008, p.18).⁴⁴ This is the first main assumption needed in order to justify normality testing as a method for determining whether stopping was non-random, and is clearly satisfied by the slot machine environment.⁴⁵

In addition, I need to assume that winnings are drawn from identical distributions in order to connect gamblers' preferences over winnings to the observed cross-sectional distributions of outcomes. However, distributions of total winnings are generally not

⁴⁴ By contrast slot machines in the United Kingdom are permitted to have payouts which follow adaptive algorithms, which is beneficial from the point of view of the casino in that they can ensure a short term profit. Compared to UK machines, US slot machines are completely independent draws, relying on the law of large numbers for the casino to obtain its expected long run profit. In addition US slot machines tend to pay out less frequently but in larger amounts when they do.

⁴⁵ A key point is whether the number of bets taken by individuals in the sample is large enough for their personal theoretical distribution of cumulative winnings to approximate normality. Appendix 2.C shows the convergence towards a normal distribution when the underlying payout distribution is shaped like the exponential distribution. For functional simplicity, I will assume from this point forward that the personal distribution of cumulative winnings conditional on wager size and number of bets taken has approximately converged to the normal distribution – however lack of full convergence does not adequately account for the pattern of stopping decisions observed in the data.

identical across gamblers for two main reasons. First, there is a drift associated with a gambler's net winnings depending on how many bets he has taken. If a single bet has an expected value of μ and a variance of σ^2 , the sum of outcomes from a sufficiently large number b of identical bets is approximated by $N(\mu b, \sigma^2 b)$. This means that in the case of negative expected value bets such as slot machine bets, as a gambler takes more bets, the expectation of his cumulative winnings is declining and its variance is increasing. In other words, there is both a negative drift and a spreading out of possible outcomes with each bet taken.

Second, the distribution of possible winnings can differ depending on the gambler's chosen wager size. This is important because assuming that the mean and variance of a single bet remain in approximately constant proportions for a given wager unit, a gambler taking 5 cent bets has a higher expected value and lower variance of winnings than a gambler betting 80 cents each time, for an identical number of bets. A simple way of modeling this potential heterogeneity in bet sizes is to categorize the distribution of winnings as a function of both number of bets taken b and wager size w , $N(w\mu b, (w\sigma)^2 b)$. I take w to be an individual gambler's average wager during the visit – the ratio of a gambler's total amount wagered over the total number of bets taken.⁴⁶

So to maintain cross-sectional normality, I need groups of gamblers who are identical across wager size as well as number of bets taken. In other words, gamblers observed in a single wagers-bets bin (w,b) should be theoretically indistinguishable from

⁴⁶ Within-person wager heterogeneity is unobserved in the data, but under the finite menu of payout distributions in a casino, survives the CLT through the Lindberg-Feller condition. Meaning with reasonable assumptions on variances of any single bet's distribution, the sum of non-identical but independent bet outcomes will still converge to a normal distribution.

a normal distribution under the EU hypothesis. However, since the data do not provide enough gamblers in each (w,b) bin to test for normality separately for each bin, consecutive bins must be combined in order to get a reasonable sample size.⁴⁷ While this will technically result in a non-normal distribution, the objective is to combine consecutive bins in such a way that simulations of winnings drawn from empirical bin frequencies of the relevant normal distributions will be nearly indistinguishable from normality. Applying the same normality test to both the simulated data and the actual data will then reveal how far from random the quitting behavior in the data were.

What does it mean to reject normality in the data? First, compared to the normal distribution in winnings implied by EU, excess kurtosis would imply an unusually high stopping frequency at some particular level of winnings, which would then be interpreted as the reference point. Figure 2.6 shows three distributions of varying kurtosis levels; the distribution on the far right has “excess kurtosis” or higher kurtosis than a normal distribution.

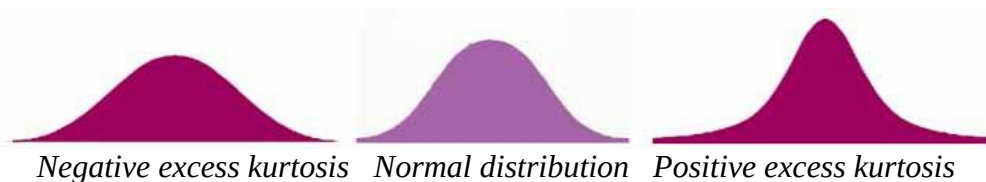


Figure 2.6: Varying kurtosis

⁴⁷ An alternative approach would sample and sum a small finite number of individual gamblers' final winnings with replacement from both the simulated normal distributions and the empirical distribution. The empirical distribution of sums should be normal in the limit under EU, and remain non-normal under PT. However, this appears to be a computationally intensive approach given the heterogeneity of normal distributions used, and would not be as visually informative about cross-sectional quitting patterns compared to combining the adjacent (w,b) bins.

Another possible way to reject normality is if the distribution has skewness. Since the normal distribution has zero skewness, excess positive skewness simply means that the distribution of winnings at quitting for a particular group of bins has a longer right tail and a shorter left tail. In other words, there is a higher overall probability of quitting above the reference point given by the peak of the distribution compared to the probability of quitting below the peak. In PT, recall that this a consequence of diminishing sensitivity in the gains and loss segments of the value function, since below the reference point gamblers are risk-loving and highly unlikely to quit, while above the reference point gamblers remain risk-averse. Figure 2.7 shows illustrations of positively (and negatively) skewed distributions.

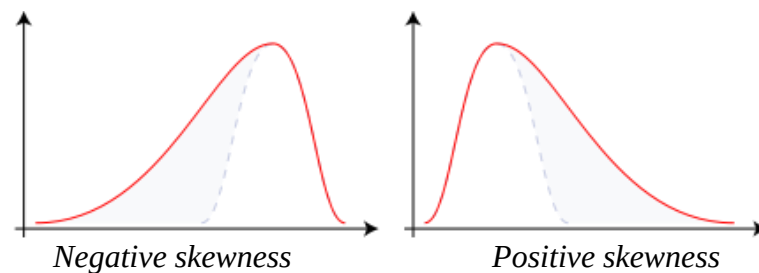


Figure 2.7: Skewness

To summarize the predictions that the two categories of quitting behavior have for the distribution of winnings in a sufficiently restricted bin range of bets taken and average wagers, EU would make the empirical distribution of winnings indistinguishable from the normal distribution, while PT would be able to reject the normal distribution. Rejection of the normal distribution alone would not be enough to accept PT as an alternative hypothesis. However if we were to then consider exactly in what way the normal distribution was rejected, PT would suggest positive excess kurtosis, and positive skewness of the winnings distributions.

2.5.1 Normality testing

Since a normal distribution has zero skewness and kurtosis of 3 regardless of the mean and variance of the distribution, a natural method to test whether data are normally distributed is to measure and compare these two higher moments in the data to those of the normal distribution. A commonly used test is by Jarque and Bera (1987) who combine skewness and kurtosis into a single omnibus test for normality. I use the small sample adjusted version of the Jarque-Bera test (henceforth AJB) provided by Urzua (1996), since the sample sizes for merged (w,b) bins are generally in the range of 20 to 60 observations per merged bin, which is insufficient for AJB to reach its asymptotic chi-squared distribution.⁴⁸ In order to obtain more precise critical values for the AJB test for sample sizes not included in the table from Urzua (1996), I simulate the 95% critical values for the range of relevant sample sizes.⁴⁹

Since w is the average wager of an individual gambler, and estimate of the true amount they prefer to wager on each bet, putting individuals into groups based on consecutive average wagers essentially groups them by an estimate of their gambling type – for example, “very small stakes” gamblers might be grouped as those who wagered an average of between one and 20 cents on each bet. “Medium stakes” gamblers might be defined as those who wagered an average of 40 to 60 cents per bet.

The total number of bets taken and average wagers for the 2393 people in the sample is shown in Figures 2.8a and 2.8b. As Figure 8a demonstrates, it was more

⁴⁸ See Poitras (2006) for a discussion of small sample properties of various normality tests.

⁴⁹ By drawing N observations 500,000 times from a generic normal distribution, and taking the 95th percentile of the AJB as the critical value. A table of the simulated critical values is in Appendix 2.A.

common for people to take few bets and increasingly less common to take larger numbers of bets. 92 percent of gamblers in the sample took 2000 bets or fewer. In Figure 2.8b, the most common average wager size was between 30 to 40 cents, with average wagers between 20 cents and 80 cents accounting for 56 percent of gamblers in the sample.

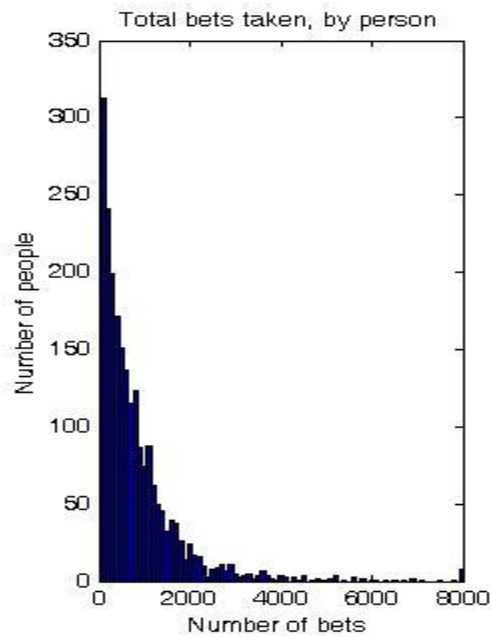


Figure 2.8a:

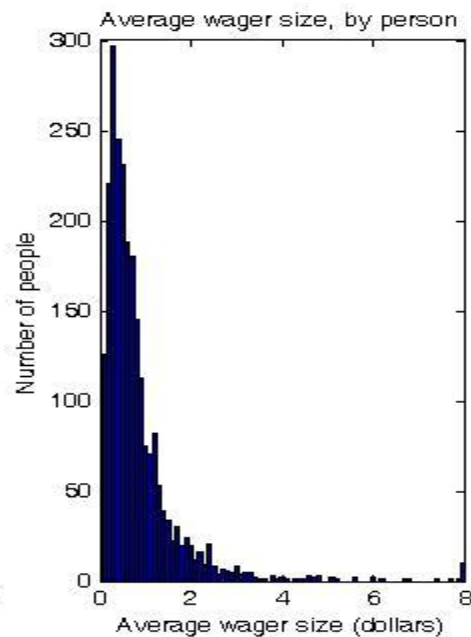


Figure 2.8b:

Column for 8000 bets (Figure 2.8a) includes gamblers taking more than 8000 bets. Column for 8 dollars (Figure 2.8b) includes gamblers with average wagers over 8 dollars.

As mentioned previously, I combine wager-bets (w, b) bins into groups to obtain a large enough number of gamblers in each group to test whether the data in those groups violate normality. In addition, since combining bins no longer yields a normal distribution even in the limit, for each group I randomize winnings outcomes according to the relevant distributions $N(w\mu b, (w\sigma)^2 b)$ based on the observed frequency of w and b in the data. This serves as a control for the natural deviations from normality that arise from combining bins. In the simulations I use $\mu = -0.05$ and $\sigma = 10.6$ which correspond to the

5 percent house advantage and the standard deviation per unit wagered respectively, suggested by Eadington (1999). The simulation results are not sensitive to small changes in these parameterizations, particularly since the AJB statistic uses centralized skewness and kurtosis measures which are normalized by the data variance.

How important is it to account for wager heterogeneity? This is a question of particular interest since a gambler's choice of wager is itself an indicator of risk preferences. If the data tracked the wager amount for each individual bet, then useful information about the dynamics of risk preferences might result – for example, I might consistently observe whether a gambler reduced his bet size when approaching certain (reference) winnings levels. However, since only the average wager for each gambler is available, the gambler's "type" as defined by whether they were a low stakes bettor, high stakes bettor or somewhere in between, is actually a confounding factor when testing for normality. This is because heterogeneity in average wagers across gamblers can produce non-normality (via heterogeneous means and variances) which would be overlooked if wagers were simply averaged across all gamblers.⁵⁰

Table 2.1 divides the sample into groups by total number of bets taken in a single visit. The columns labeled "common wager" show the simulated normality rejection rates, 95th percentile of skewness and excess kurtosis respectively, all assuming that gamblers have a common wager equal to the average wager in the population of 84 cents. The AJB test fails to reject the null hypothesis of normality about 95 percent of the time for groups with larger number of bets. Skewness and excess kurtosis at the 95th percentile are not

⁵⁰ For example two normal distributions with the same mean but different variances tend to produce excess kurtosis. Two normal distributions with very different means tend to produce a distribution with two peaks, and so on.

drastically different from those of the normal distribution. However once the same individuals are simulated according to their personal average wagers, normality is fully rejected for the exact same subgroups of gamblers (see far most right 3 columns). This shows that wager heterogeneity cannot be ignored, and therefore that the groups of gamblers in Table 2.1 must be broken into even smaller groups by wager sizes in the main empirical analysis in order to provide a valid simulated control group.

**Table 2.1: Comparison of group rejection of normality with and without wager heterogeneity
(simulations only)**

from bet	to bet	number of gamblers in group	simulated reject rate of normal distribution at 95% level (common wager = 0.84)	95th percentile of group skewness (common wager)	95th percentile of group excess kurtosis (common wager)	simulated reject rate of normal distribution at 95% level (personal wager)	95th percentile of group skewness (personal wager)	95th percentile of group excess kurtosis (personal wager)
1	60	222	0.47	0.35	1.79	1	6.61	117.99
61	130	222	0.11	0.29	0.91	1	2.01	23.91
131	210	203	0.08	0.29	0.76	1	4.84	76.25
211	300	202	0.06	0.26	0.68	1	3.30	39.03
301	410	203	0.06	0.30	0.65	1	7.96	125.72
411	540	200	0.06	0.30	0.62	1	2.32	34.88
541	700	203	0.05	0.29	0.60	1	3.44	51.54
701	870	202	0.05	0.27	0.62	1	2.30	32.47
871	1120	204	0.06	0.28	0.64	1	2.96	42.65
1121	1520	202	0.06	0.30	0.70	1	0.94	12.82
1521	2490	200	0.07	0.29	0.69	1	5.29	84.93
2491	14370	130	0.40	0.33	2.98	1	1.89	30.54

Number of bets grouped so as to have at least 200 gamblers per group. Columns 4 through 6 (common wager) are simulated using $N(0.84^*\mu_b, (0.84\sigma)^2b^*)$ with empirical b frequencies. Columns 7 through 9 (personal wager) are simulated using $N(w\mu_b, (w\sigma)^2b)$ with empirical b and w frequencies. 1000 simulations for each group.

2.6 Results

Table 2.2 shows a comparison of the simulated winnings distribution for each group of combined bins, and the AJB statistic for each corresponding group in the data. The simulated reject rate is typically near or slightly above 5 percent using the 95 percent critical values estimated in Appendix 2.A. As in Table 2.1, the reject rate is substantially higher for simulations for small numbers of bets (bets 1 to 60), which is consistent with the CLT.

The four columns on the far right of Table 2.2 compare the 95th percentile of the simulated excess skewness and kurtosis values to the values of these moments in the data. The simulated 95th percentile of skewness column shows that there is some positive skewness that results purely from combining adjacent bins – however in cases where the data rejects normality, the data tend to be skewed even farther to the right than the simulations. The story is similar for kurtosis levels. There is a natural tendency for quitting to bunch near certain winnings levels due to merging the bins, however in groups where normality is rejected in the data, the bunching exceeds the levels implied by the simulation.

Figure 2.9 shows histogram plots of the groups in Table 2.2 (smallest and largest bet categories omitted to save space), with a generic normal distribution fit overlaid in red. Groups whose winnings distributions are approximately normal should have histogram plots which roughly follow the outline of the normal distribution.

Table 2.2: Normality tests by number of bets and average wager types

from bet	to bet	from		to average wager (dollars)	number of gamblers in group	simulated reject rate of normal distribution		AJB statistic (data)	indicator for rejection of normal distribution		simulated 95th percentile of skewness		simulated 95th percentile of excess kurtosis		data excess kurtosis
		average wager (dollars)	to average wager (dollars)			at 95% critical value			in data at 95% level	skewness	skewness	data skewness	kurtosis	kurtosis	
1	60	0.21	0.40	43	0.20	0.12	0.20	9.40	1	0.83	0.17	2.89	2.15		
1	60	0.41	0.60	30	0.15	0.09	0.15	38.28	1	0.95	1.89**	2.82	3.59**		
1	60	0.61	0.80	36	0.17	0.08	0.17	1586.3	1	0.90	5.16**	3.03	28.88**		
61	130	0.21	0.40	50	0.12	0.09	0.12	41.81	1	0.62	1.41**	2.18	3.27**		
61	130	0.41	0.60	49	0.09	0.07	0.09	1.32	0	0.64	0.00	1.71	0.77		
61	130	0.61	0.80	33	0.08	0.05	0.08	66.98	1	0.78	2.16**	2.10	4.99**		
131	210	0.21	0.40	37	0.12	0.07	0.12	3.81	0	0.78	0.74	2.21	0.27		
131	210	0.41	0.60	41	0.07	0.05	0.07	4.77	0	0.67	0.37	1.65	1.41		
131	210	0.61	0.80	26	0.05	0.03	0.05	15.34	1	0.78	1.06**	1.86	2.79**		
211	300	0.21	0.40	43	0.09	0.08	0.09	29.85	1	0.67	1.45**	1.87	2.63**		
211	300	0.41	0.60	34	0.08	0.06	0.08	200.04	1	0.74	2.89**	1.88	9.61**		
211	300	0.61	0.80	26	0.07	0.04	0.07	1.95	0	0.80	-0.57	2.18	0.53		
301	410	0.21	0.40	49	0.12	0.08	0.12	1.06	0	0.68	0.18	2.03	0.59		
301	410	0.41	0.60	39	0.08	0.06	0.08	47.13	1	0.64	1.69**	1.76	3.86**		
301	410	0.61	0.80	31	0.06	0.04	0.06	85.94	1	0.66	2.33**	1.64	6.11**		
411	540	0.21	0.40	45	0.087	0.06	0.087	58.84	1	0.60	1.20**	1.83	4.77**		
411	540	0.41	0.60	36	0.061	0.04	0.061	840.02	1	0.65	4.08**	1.81	20.78**		
411	540	0.61	0.80	31	0.06	0.04	0.06	200.30	1	0.73	2.76**	1.87	10.29**		
541	700	0.21	0.40	48	0.11	0.07	0.11	42.52	1	0.66	1.30**	1.91	3.58**		
541	700	0.41	0.60	47	0.067	0.04	0.067	86.99	1	0.59	1.85**	1.56	5.21**		

Table 2.2: Normality tests, Continued

541	700	0.61	0.80	22	0.054	126.18	1	0.82	2.60**	1.94	9.44**
701	870	0.21	0.40	53	0.105	63.03	1	0.62	1.54**	2.00	4.12**
701	870	0.41	0.60	37	0.06	1.22	0	0.60	0.11	1.59	0.81
701	870	0.61	0.80	34	0.052	63.94	1	0.65	1.89**	1.65	5.10**
871	1120	0.21	0.40	47	0.116	0.99	0	0.66	0.06	2.11	-0.67
871	1120	0.41	0.60	37	0.075	178.22	1	0.69	2.35**	1.84	9.02**
871	1120	0.61	0.80	37	0.079	154.53	1	0.70	2.33**	1.84	8.26**
1121	1520	0.21	0.40	44	0.099	3.58	0	0.65	0.29	1.98	1.20
1121	1520	0.41	0.60	35	0.074	27.25	1	0.69	1.20**	1.88	3.31**
1121	1520	0.61	0.80	32	0.062	5.34	0	0.69	0.62	1.75	1.43
1521	2490	0.21	0.40	46	0.093	11.53	1	0.63	0.47	1.77	2.15
1521	2490	0.41	0.60	44	0.092	20.87	1	0.68	1.35**	1.79	1.79
1521	2490	0.61	0.80	28	0.055	12.52	1	0.75	1.15**	1.78	2.05

Simulations repeated 1000 times for each bin, drawing from $N(w_{ub}, (w\sigma)^2)$ using empirical b and w frequencies, and $\mu = -0.05$, $\sigma = 10.6$ (Eadington 1999).
 ** deviation from simulated data at the 95th percentile in direction predicted by PT. Results shown for modal wager categories only. Full table with all bins in Appendix 2.B.

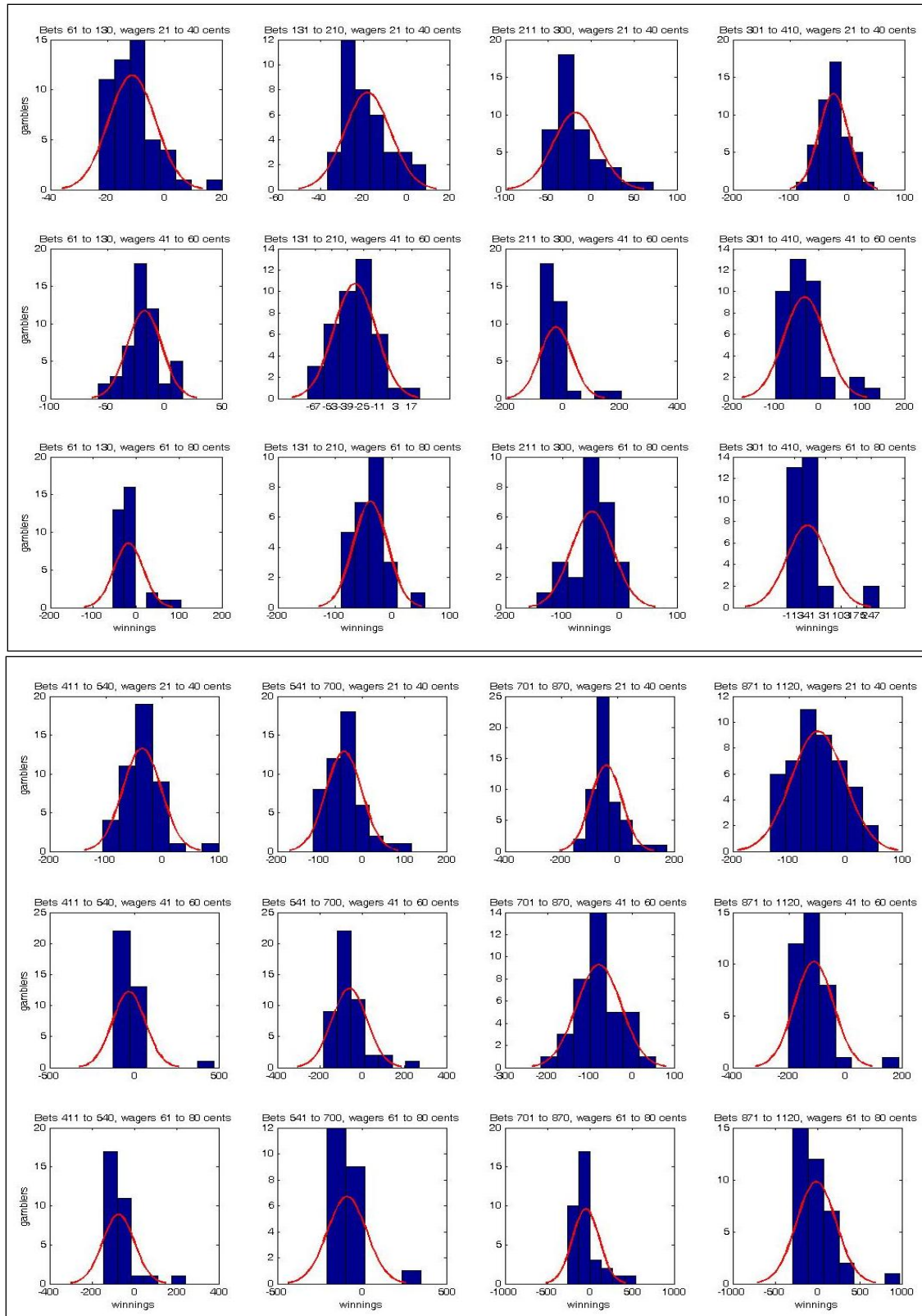


Figure 2.9: Histograms by bins

Figure 2.10 summarizes the results in Table 2.2 in a slightly different way using quantile-quantile (Q-Q) plots. In each graph, actual winnings from the data are lined up according to quantiles on the x-axis, and randomized winnings (conditioning on number of bets taken and average wager as in Table 2.2) are lined up by quantiles on the y-axis. Therefore the y-axis represents where the data points should be located under the assumption of random quitting. If the data were also generating by random quitting, then the blue points should line up on the dotted line where the theoretical and empirical quantiles match.

From left to right, each panel shows a shrinking subset of the data. The farthest left panel includes all of the data, the second panel includes only winnings between - \$1000 and \$1000, the third panel between -\$500 and \$500, and finally the panel on the right shows only -\$100 to \$100. The purpose of the four panels is only to zoom in so that the data points in relation to the dotted line can be seen more clearly. For example, in the full sample version (far left) the data appears to match the theoretical distribution, other than a few select outliers at the tails. A closer examination by zooming in however, shows that once eliminating outliers, the data actually show a consistent deviation from the theoretical distribution.

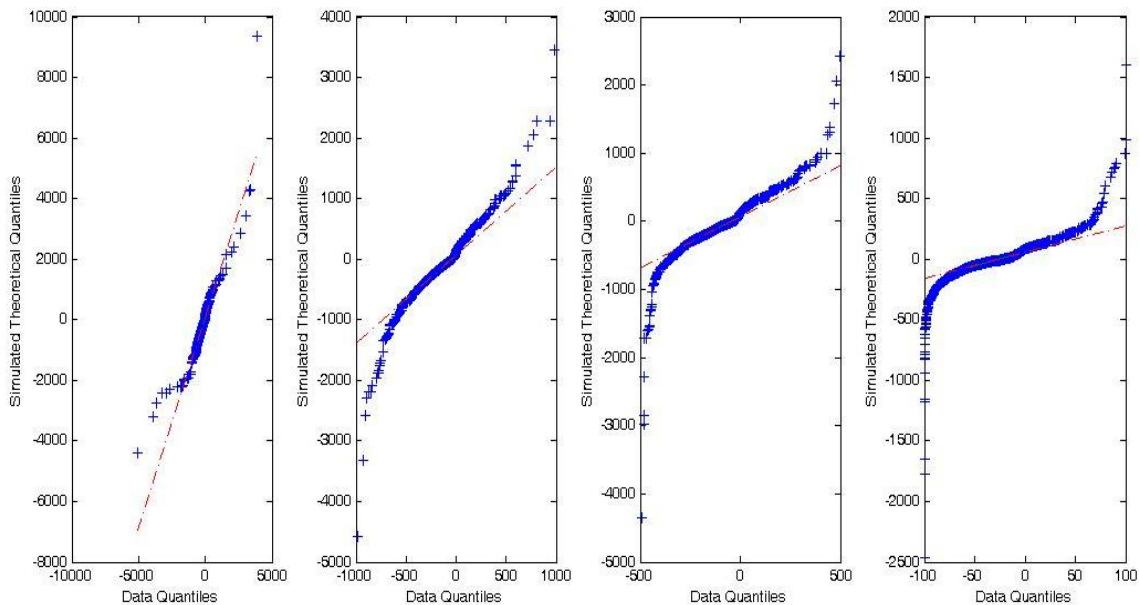


Figure 2.10: Q-Q plots

Simulated data quantiles drawn from $N(w\mu b, (w\sigma)^2 b)$ using empirical b and w frequencies, and $\mu = -0.05$, $\sigma = 10.6$ (Eadington 1999).

First, the theoretical distribution (which is a mixture of normal distributions) predicts a more spread out distribution of winnings given the numbers of bets taken and wagers made, compared to the winnings realized in the data. This can be seen by the fact that at both the winning and losing end of the distribution, the simulations predict more far reaching outliers than those present in the data.

Second, there is a clear difference between how well the data match the theoretical distribution for positive winnings compared to negative ones. Realized positive winnings in the data are consistently below the theoretical predictions for a given quantile, as evidenced by the blue data points lying above the dotted line for winnings greater than zero. This implies that gamblers tend to “run down” their winnings by re-wagering them into additional bets, more so than implied by random stopping. This is

consistent with the house money effect, in that positive winnings might be viewed as expendable bonuses to gamblers.

The patterns described in Figure 2.10 are robust to moderate changes in parameterization to the mean and variance of the payout distributions.⁵¹

2.7 Conclusions

This paper has presented empirical evidence that prospect theory preferences guide the decisions of individuals in small stakes risk-taking settings, in a slot machine environment which is simple and unchanging over time. The aggregate data demonstrate a high concentration of quitting near the break-even point, and furthermore, evidence of a reference point persists once accounting for heterogeneity in gambler enjoyment and wager choices.

Since a gambler's total winnings are simply the sum of independently distributed random variables' outcomes, the implication of EU-style random quitting via the central limit theorem, is that the distribution of bet outcomes should be approximately normal within a sufficiently narrow range of wager sizes and bets taken. The analysis showed that people are in fact disproportionately prone to quitting within a narrow band of winnings, which can be interpreted as containing the PT reference point. This tendency resulted in higher-than-normal kurtosis levels in most of the realized net payoff

⁵¹ Drastically decreasing the mean of the single bet distribution alters the Q-Q plot to reflect stronger skewness, but still do not improve the fit of the data to the straight line. The patterns in Figure 2.10 are robust to almost any change in the standard deviation of a single bet distribution (including halving or doubling the value of σ). I interpret this as an indication that the data are difficult to reconcile with random quitting, regardless of the assumptions about the mean and variance of the outcome of a single bet.

distributions, suggesting that gamblers valued their winnings according to prospect theory preferences and subscribed to the house money and break-even effects.

Also, based on the distributions of realized winnings, gamblers were more likely to quit when above the reference point suggested by the peak of the distribution, compared to quitting below it. In other words, the final winnings distributions had positive skewness, implying that the break-even effect tended to exceed the house money effect, ie. that either the PT value function's gain or loss segment exhibited diminishing sensitivity.

A shortcoming of the analysis is the use of average wager by individual gamblers as an estimate for a gambler's "type". Since gamblers may potentially change the size of their wager on each bet they take, using the ratio of money spent gambling to total bets taken is only a rough proxy to summarize their gambling behavior. The use of average wager also does not account for changes in how much a gambler decides to wager after observing the outcome of their bets. For example, a gambler who has recently won may increase the size of his wager in response to the win, and one who has lost may subsequently decrease his wager. The analysis in this paper does not account for these possibilities since bet-by-bet wagers and outcomes are unobserved, however simulations of heterogeneous bets suggest that varying the wager amount within a small range should not affect predictions on gambler outcomes.

In modeling the slot machines I also assumed that each of them had identical payoff distributions with a fixed per-unit-wagered mean and standard deviation. Although the exact value of the mean and standard deviation does not impact the results,

the analysis did not attempt to explain how gamblers select which machine to play, or allow for the possibility of heterogeneous payoff distributions other than conditioning on wager amount. In the case of heterogeneous wagers, I assumed that the mean and standard deviation would remain proportional to each other. It might be that certain machines yield better return distributions than others in exchange for being a new or relatively unpopular game. These and other possibilities associated with machine heterogeneity are not accounted for in this analysis, although one would not expect significant arbitrage opportunities to exist in equilibrium based on machine choice or wager size.

A final caveat is that the analysis cannot technically distinguish between PT and a variation of EU in which the utility of gambling is dependent on the level of winnings. For example, a form of preferences where gamblers find continuing to bet intrinsically more valuable when they are ahead due to feeling lucky, and intrinsically more valuable when they are behind because they feel they are due to win something. With this dataset it is not possible to untangle the potential interactions between gamblers' valuation of wealth and their valuation of making bets.

In spite of these shortcomings, the statistical tests on the data clearly show violations of normality in the winnings of small groups of gamblers who were sufficiently similar in wagering and betting. Due to the independent nature of slot machine bet outcomes, this strongly suggests that there was more reasoning behind gamblers' quitting behavior than pure randomness. A prospect theory explanation where

gamblers in the same group have a common reference point is a plausible explanation which is also consistent with the specific violations of normality in the data.

Finally, this study is one of several field environments where the implications of prospect theory and expected utility have been compared and tested, adding to the literature of field settings where PT offers the most plausible explanation of behavior. Future work should consider similarly stochastic environments where the expected value of the bet is not always negative, where variances of bet outcomes are perhaps more controlled, and where populations of participants vary.

Chapter 2 has been submitted for publication. The dissertation author was the primary investigator and author of this paper.

Appendix 2.A:

Table 2.A.1: Simulated 95% critical values for adjusted Jarque-Bera test
(Urzua 1996)

N	CV	N	CV	N	CV
10	7.357	27	6.816	44	6.601
11	7.341	28	6.738	45	6.607
12	7.199	29	6.819	46	6.627
13	7.186	30	6.768	47	6.578
14	7.114	31	6.764	48	6.575
15	7.102	32	6.749	49	6.524
16	7.033	33	6.680	50	6.571
17	7.019	34	6.654	51	6.491
18	6.909	35	6.641	52	6.573
19	6.989	36	6.644	53	6.530
20	6.935	37	6.700	54	6.543
21	6.935	38	6.685	55	6.491
22	6.925	39	6.683	56	6.515
23	6.833	40	6.683	57	6.500
24	6.790	41	6.614	58	6.486
25	6.802	42	6.621	59	6.477
26	6.853	43	6.613	60	6.447

Simulated using 500,000 replications of the Adjusted Jarque-Bera statistic derived by Urzua (1996), with sample drawn from $N(0,1)$ distribution. Note that critical values tend to be slightly lower than the published values, possibly due to higher number of replications us

Appendix 2.B: Results for all grouped bins

Table 2.B.1: Normality tests by number of bets and average wager types

from bet	to bet	from average wager (dollars)	to average wager (dollars)	number of gamblers in group	simulated reject rate of normal distribution at 95% critical value	adjusted JB statistic	rejection of normal distribution in data at 95% level	simulated 95th percentile of skewness	data skewness	simulated 95th percentile of excess kurtosis	data excess kurtosis
1	60	0.01	0.20	14	0.15	0.35	0	1.27	-0.03	4.01	-0.68
1	60	0.21	0.40	43	0.20	9.40	1	0.83	0.17	2.89	2.15
1	60	0.41	0.60	30	0.15	38.28	1	0.95	1.89	2.82	3.59
1	60	0.61	0.80	36	0.17	1586.31	1	0.90	5.16	3.03	28.88
1	60	0.81	1.00	24	0.13	75.55	1	0.91	2.03	2.93	6.94
1	60	1.01	1.50	31	0.16	19.06	1	0.95	1.40	2.99	2.31
1	60	1.51	2.50	27	0.19	174.35	1	1.09	2.61	3.53	10.33
1	60	2.51	20.33	17	0.57	258.12	1	2.68	3.81	11.69	15.41
61	130	0.01	0.20	24	0.12	30.19	1	0.94	1.73	2.77	3.76
61	130	0.21	0.40	50	0.12	41.81	1	0.62	1.41	2.18	3.27
61	130	0.41	0.60	49	0.09	1.32	0	0.64	0.00	1.71	0.77
61	130	0.61	0.80	33	0.08	66.98	1	0.78	2.16	2.10	4.99
61	130	0.81	1.00	19	0.06	0.02	0	0.84	0.00	2.17	-0.14
61	130	1.01	1.50	24	0.08	512.75	1	0.82	4.26	2.28	19.06
61	130	1.51	2.50	15	0.08	16.97	1	1.10	1.86	2.55	2.91
61	130	2.51	20.33	8	-	-	-	-	-	-	-
131	210	0.01	0.20	31	0.21	21.47	1	0.96	1.47	3.37	2.50
131	210	0.21	0.40	37	0.12	3.81	0	0.78	0.74	2.21	0.27
131	210	0.41	0.60	41	0.07	4.77	0	0.67	0.37	1.65	1.41
131	210	0.61	0.80	26	0.05	15.34	1	0.78	1.06	1.86	2.79
131	210	0.81	1.00	17	0.06	3.84	0	0.94	0.81	2.23	1.37
131	210	1.01	1.50	22	0.06	47.94	1	0.85	2.16	2.17	5.10
131	210	1.51	2.50	19	0.08	3.38	0	0.93	0.85	2.56	0.86

Table 2.B.1: Normality tests, Continued

131	210	2.51	20.33	10	0.25	20.38	1	1.85	1.95	5.55	4.68
211	300	0.01	0.20	32	0.24	9.57	1	1.02	0.98	3.66	1.60
211	300	0.21	0.40	43	0.09	29.85	1	0.67	1.45	1.87	2.63
211	300	0.41	0.60	34	0.08	200.04	1	0.74	2.89	1.88	9.61
211	300	0.61	0.80	26	0.07	1.95	0	0.80	-0.57	2.18	0.53
211	300	0.81	1.00	11	0.07	1.60	0	1.14	-0.57	2.89	-1.18
211	300	1.01	1.50	28	0.06	181.34	1	0.74	2.76	1.81	10.23
211	300	1.51	2.50	23	0.08	51.32	1	0.91	1.95	2.33	5.52
211	300	2.51	20.33	5	-	-	-	-	-	-	-
301	410	0.01	0.20	25	0.24	40.75	1	1.16	0.38	4.14	5.71
301	410	0.21	0.40	49	0.12	1.06	0	0.68	0.18	2.03	0.59
301	410	0.41	0.60	39	0.08	47.13	1	0.64	1.69	1.76	3.86
301	410	0.61	0.80	31	0.06	85.94	1	0.66	2.33	1.64	6.11
301	410	0.81	1.00	17	0.053	132.54	1	0.90	3.02	2.10	10.76
301	410	1.01	1.50	24	0.056	1.69	0	0.70	-0.38	1.84	-0.93
301	410	1.51	2.50	11	0.06	1.36	0	1.13	-0.33	2.71	-1.34
301	410	2.51	20.33	7	-	-	-	-	-	-	-
411	540	0.01	0.20	22	0.26	178.02	1	1.14	2.95	4.50	11.35
411	540	0.21	0.40	45	0.087	58.84	1	0.60	1.20	1.83	4.77
411	540	0.41	0.60	36	0.061	840.02	1	0.65	4.08	1.81	20.78
411	540	0.61	0.80	31	0.06	200.30	1	0.73	2.76	1.87	10.29
411	540	0.81	1.00	19	0.052	2.48	0	0.88	0.82	2.04	0.20
411	540	1.01	1.50	26	0.081	2.43	0	0.82	0.41	2.28	1.12
411	540	1.51	2.50	16	0.07	12.98	1	0.92	1.00	2.45	3.43
411	540	2.51	20.33	5	0.051	0.89	0	1.58	0.04	3.68	-1.88
541	700	0.01	0.20	24	0.199	0.35	0	1.03	-0.22	3.88	0.32
541	700	0.21	0.40	48	0.11	42.52	1	0.66	1.30	1.91	3.58
541	700	0.41	0.60	47	0.067	86.99	1	0.59	1.85	1.56	5.21
541	700	0.61	0.80	22	0.054	126.18	1	0.82	2.60	1.94	9.44
541	700	0.81	1.00	17	0.044	33.70	1	0.95	1.99	1.98	4.83
541	700	1.01	1.50	29	0.071	64.50	1	0.78	1.97	1.91	5.60
541	700	1.51	2.50	13	0.065	36.38	1	0.99	2.41	2.62	5.47

Appendix 2.C: Within-person convergence rates

This section discusses the speed at which cumulative winnings of an individual gambler converge to the normal distribution under the assumption of an exponential underlying slot machine payout distribution. I consider both within person heterogeneity in wager choice by gambler, as well as homogenous wagers – however, there is little effect of within-person heterogeneity in wagers on convergence.

Recall in the analysis in the main body of the paper, a normal distribution of the form $N(w\mu_b, (w\sigma)^2b)$ was assumed in order to generate comparison samples for bins sorted by average wager and number of bets taken. Therefore, one would hope that the normal distribution is approximated reasonably rapidly.

The exponential distribution is used as a model of the underlying payout distribution because it has certain features which mimic the intuition of a slot machine gamble. First, there is a limit to how much a gambler loses (ie. the wager), which can be obtained by shifting the exponential distribution left by the wager amount so that the new support of the distribution is $[-w, \infty)$. Second, the gambler is less and less likely to receive a payout conditional on the amount paid out by the machine, consistent with the downward slope of the distribution function. Finally, there is a small possibility of a large payout or “jackpot”, as in the exponential distribution, although this amount will be finite in practice. Medium or moderately large sized “jackpots” are also unlikely to be continuously distributed as implied by the exponential distribution. In these respects, using the exponential distribution as a model for underlying slot machine payouts might

be viewed as a type of worst-case-scenario simulation since the underlying distribution is very different from the supposed asymptotic one.

Table 2.C.1 below shows the convergence of the sum of exponential random variables to the normal distribution at the 90th and 95th percentiles as the number of bets taken increases. Note that the exponential distribution itself has skewness of 2 and kurtosis of 6, and the normal distribution has skewness of 0 and kurtosis of 3. Table 2.C.1 shows that it does not take many bets for the sums of exponentially distributed random variables to reduce the original skewness and kurtosis drastically. At the same time, complete convergence to the normal distribution occurs more slowly as the number of bets increases.

Table 2.C.1: Convergence of cumulative winnings drawn from exponential distribution

	iid bets: $w = .84$				non-identical (independent, uniformly distributed) bets			
	skewness (90th percentile)	kurtosis (90th percentile)	skewness (95th percentile)	kurtosis (95th percentile)	skewness (90th percentile)	kurtosis (90th percentile)	skewness (95th percentile)	kurtosis (95th percentile)
$b = 30$	0.52	3.68	0.58	3.87	0.67	4	0.72	4.26
$b = 50$	0.5	3.68	0.56	3.87	0.55	3.71	0.6	3.9
$b = 100$	0.42	3.49	0.45	3.69	0.43	3.56	0.47	3.74
$b = 200$	0.34	3.4	0.38	3.51	0.34	3.39	0.38	3.56
$b = 300$	0.31	3.4	0.34	3.5	0.31	3.37	0.36	3.5
$b = 400$	0.27	3.37	0.31	3.49	0.29	3.35	0.34	3.5
$b = 500$	0.26	3.32	0.3	3.44	0.27	3.33	0.31	3.4
$b = 600$	0.26	3.29	0.29	3.44	0.24	3.28	0.29	3.39

Simulations of the bins in Table 2.2 summing outcomes from an exponentially-shaped underlying payout distribution as in Table 2.C.1, show that slow convergence of this form is unable to account for the drastic deviations from normality in the Table 2.2 bins for actual data.

References:

- Barberis, N. "A Model of Casino Gambling", March 2009.
- Barberis, N., and Xiong, W. "What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation" NBER Working Paper No. 12397, July 2006.
- Camerer, C. "Prospect Theory in the Wild: Evidence from the Field", Caltech Social Science Working Paper 1037, May 1998.
- Camerer, C., Babcock, L., Loewenstein, G., and Thaler, R. "Labor Supply of New York City Cabdrivers: One Day at a Time" *The Quarterly Journal of Economics*, May 1997, p. 407 - 441.
- Conlisk, J. "The Utility of Gambling" *Journal of Risk and Uncertainty*, 6:255-275 (1993).
- Crawford, V.P., and Meng, J.J. "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational Expectations Targets for Hours and Income" UCSD Working Paper, 2008-03, July 2008.
- Deb, P., and Sefton, M. "The distribution of a Lagrange multiplier test of normality" *Economics Letters* 51, (1996) p. 123 – 130.
- Eadington, W.R. "The Economics of Casino Gambling" *The Journal of Economic Perspectives*, Vol. 13, No. 3 (Summer 1999) pp. 173-192.
- Farber, H.S., "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers" *Journal of Political Economy*, Vol. 113, No. 1, 2005, p. 46 – 82.
- Farber, H.S., "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers" *American Economic Review*, 98:3, June 2008, p.1069 - 1082.
- Genesove, D., and Mayer, C. "Loss Aversion and Seller Behavior: Evidence from the Housing Market" *Quarterly Journal of Economics*, November 2001, p.1233 - 1260.
- Jarque, C.M., Bera, A.K. "A Test for Normality of Observations and Regression Residuals" *International Statistical Review*, Vol. 55, No.2 (Aug 1987), pp.163 – 172.
- Kahneman D., and A. Tversky. "Prospect Theory: An Analysis of Decision Under Risk" *Econometrica*, Vol. 47, No. 2, March 1979, p. 263-291.
- Kiefer, N.M., "Economic Duration Data and Hazard Functions" *Journal of Economic Literature*, Vol. XXVI, June 1988, p. 646-679.

Koszegi, B., and Rabin, M. "A Model of Reference-Dependent Preferences" *The Quarterly Journal of Economics*, Vol. CXXI, Issue 4, November 2006, p. 1133 – 1165.

Koszegi, B., and Rabin, M. "Reference-Dependent Risk Attitudes" *American Economic Review*, Vol. 97, No. 4, September 2007, p. 1047 – 1073.

List, J.A., "Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace", *Econometrica*, Vol. 72, No. 2, March 2004, p. 615 – 625.

Odean, T., "Are Investors Reluctant to Realize Their Losses?" *The Journal of Finance*, Vol. LIII, No. 5, October 1998, p.1775 – 1798.

Poitras, G., "More of the correct use of omnibus tests for normality" *Economics Letters* 90 (2006), p.304 – 309.

Post, T., van den Assem, M.J., Baltussen, G., and Thaler, R.H. "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show" *American Economic Review*, Vol. 98, No. 1, 2008, p. 38-71.

Rabin, M. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem" *Econometrica*, Vol. 68, No.5, Sept. 2000, p.1281-1292.

Rabin, M., and R.H. Thaler. "Anomalies: Risk Aversion" *Journal of Economic Perspectives*. Vol. 15, No.1, Winter 2001, p. 219 - 232.

Segal, U. and Spivak, A. "First Order versus Second Order Risk Aversion" *Journal of Economic Theory*, 51, 111-125 (1990), 111-125.

Thaler, R. and E. Johnson. "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice" *Management Science*, Vol. 36, No. 6, June 1990, p. 643-660.

Urzua, C.M., "On the correct use of omnibus tests for normality" *Economics Letters* 53 (1996), p.247-251.

Zangeneh, M., Blaszczyński, A., Turner, N.E., In the Pursuit of Winning: Problem Gambling Theory, Research and Treatment, "Chapter 2: Explaining Why People Gamble" by Walker, M., Schellink, T. and Anjoul, F.

Wooldridge, J.M. Econometric Analysis of Cross Section and Panel Data, MIT Press, 2002.

Chapter 3: Impatiently Saving for Leisure: Layaway as a Loose Commitment Savings Contract

3.1 Introduction

How do individuals choose the features of their commitment device? In this paper, I analyze a unique data set of layaway contracts in which consumers choose the intensity of their commitment to purchase an item through their choice of contract length. The specific rules of the layaway contracts in the dataset provide a strong case that consumers are using them as a savings commitment mechanism. Controlling for the price of the item ordered, consumers tended to select longer contracts or lower monthly commitment for household items and clothing, while selecting shorter contracts or higher monthly commitment for entertainment items such as video game systems, luxury goods, and multimedia items.

Empirical studies have documented the demand for commitment devices as solutions to self-control or intra-household control problems. Ashraf, Karlan and Yin (2006) establish a relationship between present-bias in time-preference experiments and participation in a savings commitment, in a field experiment conducted on bank customers in the Philippines. They find that particularly among women, higher degrees of present-bias are associated with higher likelihood of participation in a voluntary savings commitment program offered to them.⁵²

⁵² Ashraf et al. hypothesize that the gender gap is due to a cultural tradition in the Philippines that women are responsible for household finances.

Gine, Karlan and Zinman (2009) examine take up and success rates of a smoking cessation commitment experiment, also in the Philippines. Smokers deposited money into a savings account that would be forfeited if a future urine test for nicotine was positive. They found that individuals who were offered the commitment program were 3.3% to 5.8% more likely to pass the urine test 6 months later.

Much of the empirical work on commitment devices has focused on take up and success rates, but few papers have studied how chosen features of the commitment contract may vary with the individual's commitment objective.⁵³ Consumers in this layaway program choose a single feature of their layaway plan - the length of the contract, allowing observation of the chosen tradeoff between monthly sacrifice of liquidity and the time until delivery for various types of items.

I interpret a greater amount of money automatically deducted from a customer's bank account on a monthly basis as a greater sacrifice, and hence a more serious intent of commitment towards purchasing the item. That is, for an item of a given price, choosing a shorter contract entails a greater monthly commitment of resources away from very short-run affordable indulgences the consumer may be susceptible to, while hastening the arrival of the desired item. A longer contract requires a lower monthly commitment but spreads the payments over a longer time horizon.

An additional contribution of this paper is to document an example of a formal commitment arrangement that has arisen naturally in the marketplace. Since customers must pay a commission for the layaway service - a price premium above what it would

⁵³ Another relevant website is stickK.com, which allows users to set up customized commitment contracts online for any goal, such as weight loss, exercise, quitting smoking, etc.

cost them to save the money on their own in their bank accounts or at home - this data provides evidence of a market demand for commitment, and a firm which in turn makes its profits by supplying it.

Finally, another key related paper is by Oster and Scott Morton (2005) which examines the relationship between newsstand and subscription prices for 'leisure' magazines and 'investment' magazines. They find that investment magazines (those which have reading costs but yield future benefits) have higher ratios of subscription to newsstand prices than leisure magazines (those which provide immediate payoff but may have future costs). They attribute this theoretically to present-biased yet sophisticated consumers' willingness to pay a premium for 'investment' magazine subscriptions as a type of commitment device.

The remainder of the paper is organized as follows: Section 3.2 discusses the concept of layaway plans and the rules of the online layaway contractor; Section 3.3 discusses models of self-control which have inspired the subsequent literature on commitment devices; Section 3.4 covers the data, how items were categorized, and summary statistics; Section 3.5 contains the regression results; Section 3.6 concludes.

3.2 Layaway contracts

3.2.1 Historical context

A layaway contract is an arrangement, typically between a retailer and a consumer, in which the consumer agrees to make a series of payments toward the purchase of an item. The customer receives the item only *after* the price has been paid in full, distinguishing a layaway plan from an installment plan where the item is delivered

immediately and financed on credit. In layaway, the consumer forgoes the interest that could have been earned on the payments, and furthermore is typically charged either a fixed or percentage fee for the layaway service. In the event of cancellation, the customer's money is refunded short of a cancellation fee.

Layaway contracts are often associated with the Great Depression, a historical period before consumer credit became widely available, and when disposable income was scarce. These traditional layaway contracts typically involved the customer making physical payments at the store on a regular or ad-hoc basis, while the retailer saved the item for the customer in a storage room at the original price at the time of the down payment. As credit became more widely available to consumers, installment plans gained favor instead.

In the last couple of years however, layaway experienced a degree of resurgence. A New York Times article from 2008 hints at a self-control motive. "...it seems the suddenly more attractive feature of the layaway is the manner in which it enforces spending discipline....Layaway is 'smart' because (it) makes it harder to behave foolishly."⁵⁴

3.2.2 Online layaway rules

The online layaway contract works as follows. Consumers enter through the third party layaway contractor's website where if they are first time customers, they enter their identification and bank account information, becoming members of the site. They then shop on the site via a large selection of online merchants that are affiliated with the third

⁵⁴ <http://www.nytimes.com/2008/11/30/magazine/30wwln-consumed-t.html>

party contractor. Once they have decided to put an item on layaway, they select their payment schedule using the contractor's calculator shown below in Figure 3.1.

As the calculator demonstrates, the estimated ship date, number of payments (with adjustable radio buttons), monthly payment amounts, and total amount to be charged to the customer, are shown clearly to the customer. In addition, customers may further customize the exact dates of each payment if desired.

The online layaway plans I examine differ from traditional layaway in two important respects, stemming from the fact that the layaway contractor is a separate entity from the actual supplier or merchant of the ordered item: first, the availability of the item is conditional on the merchant's inventory at the time that the item is fully paid off, and therefore is not guaranteed as in traditional layaway; second, the price of the item is subject to increases dictated by the merchant while the layaway contract is in progress, and therefore is also not guaranteed at the time of contract. This eliminates two traditional reasons why a credit or income constrained consumer might find layaway more appealing than saving on their own.

Your Payments

Estimated Shipping Date: March 20, 2009

Monthly Payment

+

payments

monthly

5 @ \$20.40

-

Use the +/- buttons to adjust number of payments

Payment#	1	2	3	4	5
Amount	\$20.40	\$20.40	\$20.40	\$20.40	\$20.40
Due date	Today	12/13/08	1/13/09	2/13/09	3/13/09
Adjust due date		↑ ↓	↑ ↓	↑ ↓	↑ ↓

Use up/down buttons to adjust the due dates of your individual payments

X Hide/show payment schedule

Total: **\$102.00**

(Replica of calculator)

SUBMIT LAYAWAY

Figure 3.1: Replica of online layaway calculator

The layaway contractor charges a commission of just under 2% of the purchase price of the item.⁵⁵ The customer pays this commission concurrently with the payments for the item itself (including shipping charges and tax). For contracts which are fulfilled, the customer thus pays a total price of 1.02 times the item's retail price, shipping and tax, plus any opportunity cost (such as interest and liquidity value) of leaving the monthly payments in their bank account over the course of the contract. Thus the price premium

⁵⁵ In select cases (just over 10% of the sample), the commission was waived. This was either due to specific promotions offered by the contractor, or if the customer was a premium member. Premium members paid \$9 a month for a default membership length of a year, but cancelable on a monthly basis. In addition to waiving of the layaway commission, the premium membership involved opportunities to improve credit score history and fraud protection. In the event of customer-initiated cancellations, premium members were still responsible for cancellation fees. Slightly less than 9% of all layaway orders in the sample were attributed to a customer who at some point in time had been a premium member. Appendix B contains regression results that include a dummy variable for whether commission was charged on the layaway item. The results accounting for commissions are not quantitatively or qualitatively different than the main results.

for the layaway service consists of the commission as well as the opportunity cost due to constrained liquidity. In the event of canceling the contract, the customer pays a penalty of either \$25 or 10% of the price of the item, whichever is less.⁵⁶ Note that in contrast to the field experiments on commitment contracts, the penalty for changing one's mind is only a fraction of the total funds that the customer places into the program.⁵⁷

While macroeconomic factors such as availability of credit may partially explain why layaway plans have recently become more popular as an alternative payment mechanism, it does not explain why consumers do not simply save their money at home or in their bank accounts instead. This is particularly striking given that a bank account is a requirement to sign up for the layaway contract and already has some structural barriers to spending compared to storing cash at home.⁵⁸

Bryan, Karlan, and Nelson (2010) define a commitment device as “an arrangement entered into by an agent which restricts his or her future choice set by making certain choices more expensive....” Furthermore, the agent must be willing to pay extra for this without receiving other benefits for the payment. The online layaway program fits this criterion in that the same savings objective could be attained on their own at a lower cost – however the key benefit which the consumer is willing to pay for is

⁵⁶ In the event that the item's price decreases while the layaway contract is still in progress, the difference is refunded to the customer. In the case that the item price increases while the contract is in progress, the customer is offered the option of paying the difference or receiving a full refund. The contractor also typically tries to find the customer a comparable item that is available at the original price. Price decreases are more common than price increases.

⁵⁷ Although the rule of requiring complete forfeit of funds in the event of cancellation as in Ashraf et al. (2006) and Gine et al. (2009) is feasible in field experiments, in a purely field setting such a policy is likely to result in drastically reduced participation due to consumer risk-aversion under slight uncertainties about future preferences for items. It may be reasonable to view the layaway program as having reached a broader population since the price and thus strength of the commitment is lower.

⁵⁸ Assuming that going to the ATM requires effort, a bank account can also be seen as a commitment device. The availability of debit cards however, may reduce this barrier.

the automatic withdrawal of payments from their bank account, which restricts their month to month spending because less money is then at their disposal in the relative convenience of the bank account. A second criteria specified by Bryan et al (2010) is that the arrangement must not have a strategic purpose with respect to other individuals.

Intra-household bargaining is a motive which cannot be either ruled out or confirmed in this data set. In other words, I cannot distinguish between potentially strategic motives towards the 'self' and strategic motives towards others. In particular, Anderson and Baland (2002) have suggested that intra-household disagreement on spending can be a factor in the demand for commitment savings devices (ROSCAS) in developing countries. Although the data in this paper is not from a developing country, household bargaining may still be a significant motive.⁵⁹

3.3 Theoretical motivation

Recent literature on commitment devices and contracts focuses on their usefulness in solving self-control problems. The introduction of self-control concepts to economics is originally by Strotz (1955), who proposed time-inconsistency as an explanation for individuals engaging in costly restriction of their own future behavior. Laibson (1997) further proposed a quasi-hyperbolic discounting model, which has subsequently been adapted to explain self-control problems in a variety of settings (for example, Dellavigna and Malmendier (2006) in gym membership). Thaler and Shefrin (1981), and Fudenberg

⁵⁹ The deduction of payments from the bank account appears only in the name of the contractor and does not reveal the name of the store or item purchased. Therefore a customer who has a joint account with a spouse and wants to spend money from the account without revealing to their partner specifically what they had purchased, would have a strategic motive to participate in this layaway program. In addition, one can imagine scenarios where two married people have different preferences over consumption, and use a layaway contract as a costly agreement for spending allocations.

and Levine (2006) propose “dual-self” models, in which the objectives of the short-run self are in conflict with those of the long-run self. Finally, Gul and Pesendorfer (2001) model individuals as having a fundamental dislike of temptation.⁶⁰

Although each model takes a slightly different approach, the common result is that an agent who is aware, or ‘sophisticated’ about future temptations he or she will encounter, will attempt to restrict his or her future behavior in the present. The self-control rationale for a commitment device such as a layaway plan might be as follows:

Alice wants to purchase a plasma television which is currently ‘unaffordable’ to her.⁶¹ She could save the money in her checking or savings account over the necessary months, and independently accumulate the amount needed to purchase the desired item in full. However, if Alice has a self-control problem, she will tend to spend the disposable income in her bank account on various already-affordable items (ex. sweets, drinks, clothing, or whatever her short term temptation may be) while waiting for the money for the TV to accumulate. As a result, saving enough to afford the TV will take Alice longer than originally planned. Depending on the severity of her self-control problem, one can imagine scenarios where purchase of the TV may never come into fruition as the smaller consumption desires consistently chip away at Alice’s accumulation of disposable income.⁶²

⁶⁰ Bryan, Karlan and Nelson (2010) provide a comprehensive summary of the literature on commitment devices as well as a summary of theoretical motivations.

⁶¹ Note that “unaffordable” may be relative to a customer’s disposable income minus other monthly necessities, not necessarily absolutely unaffordable with respect to the total cash available.

⁶² A story that reflects a spousal bargaining motive might be as follows: Alice wants to purchase a plasma television which may or may not be affordable for her and her spouse Charlie. Charlie does not want the television, but would rather purchase fishing equipment. Since the fishing equipment is less expensive than the television, Alice agrees to let Charlie purchase the equipment in full, if he promises that from that point

Foreseeing this problem, Alice signs up for an online layaway contract, paying a fee in order for the contractor to automatically deduct \$X from her checking account each month so that this amount \$X will not be spent on short-term indulgences. Alice has essentially hired the contractor to be the guardian of her savings, making it such that she is unable to spend that \$X on sweets and drinks even if she wanted to (short of canceling the arrangement at a penalty). Alice trades off how much she wants to sacrifice on a monthly basis against how soon she wants to get the TV. Assuming that Alice does not cancel, the TV is delivered to her at the planned date.

3.4 Data

The data consists of the layaway contractor's new customers from June 2008 to June 2009, subsequently following all of the transactions for these customers through November 2009. There are a total of 1668 customers in the dataset. I observe the length of contract chosen by the customer for each item ordered, the item's name, each payment that was successfully made, and whether the order was canceled or fulfilled. After the initial contract was arranged, customers were in fact allowed to alter their payment schedules at no additional cost – therefore the arrangement is best thought of as a ‘loose commitment’.⁶³

Items ordered were first broadly categorized using keyword searches for common terms (such as “iPod”, “Xbox”, “14k”, etc.). Categories were created with the goal of

forward they will save for the TV. In order to ensure that Charlie does not back out of the agreement by attempting to buy more fishing equipment next month, they sign up for a layaway contract to remove money for the television periodically from their joint bank account.

⁶³ I observe only the initial contract length and the most current version of the schedule of payments. Note that this costless adjustment implies that fully attentive customers should never fail to make a payment due to insufficient funds or face overdraft charges from their bank. However, such cases occasionally do exist in the data.

accommodating both the frequency of occurrence within the data set, as well as the potential to capture differentiated contract choice between groups of items. For example, electronics is a common purchase within the data set. The broad category of electronics was then separated into a computing and non-computing category. The reasoning is that computing equipment might be considered more of a necessity to consumers than other types of electronics goods such as HD televisions or Blu-ray players.⁶⁴

When search terms failed to categorize an item, it was done manually. In addition, if the nature of the item was not obvious based on the name, additional information was obtained via web search. A more detailed description of the items in each category is shown in Appendix A. Ninety-three items in the sample were unidentifiable based on the name of the product, and were dropped from the analysis, leaving a total of 3000 items in the sample.

Categorizing items based purely by product name and type will clearly result in some customer-specific error, since I am implicitly making some assumptions regarding how an item is used. A guitar for example, may be a leisure item for an individual who is an accountant for a living, but is more of a necessity for someone who is a musician for a living.⁶⁵

⁶⁴ Categorizing items relatively narrowly turns out to be necessary for the results since significant differences (relative to the comparison group of household items) in contract length exist within the broad categories, that can be pinpointed by disaggregation. Differentiating ‘sports tickets’ from ‘other recreation,’ and ‘jewelry’ from ‘other luxury items’ are examples where aggregating across these finer categories makes both groups statistically indistinguishable from household items on average. The results tend to hold when aggregating ‘toys’ and ‘other electronics’ into a general multimedia category since both have significantly shorter contract lengths compared to household items.

⁶⁵ In this case, since musicians are a small percentage of the general population, I categorize guitars as a recreational item.

Furthermore, for several categories and items it can be unclear whether a product is fundamentally leisurely in nature or necessary. For example, a popular type of item in the sample is automotive parts. On the one hand, automotive parts are intuitively a necessity – people need working cars in order to get to their jobs. However, repairing and modifying/enhancing cars is also a hobby for some people, which would not be as necessary. Since the interpretation of this category of good is ambiguous, I leave it in its own category.⁶⁶

Summary statistics for the frequency, average price, and average contract length of various types of items ordered in the data set is shown in Table 1 below. Contracts were officially allowed to vary between three months and thirteen months long.

Table 3.1: Summary statistics

	<i>percentage</i>	<i>average price</i>	<i>average contract length</i>
toys (including video games, mp3 players)	18.4%	255.25	4.82
gift cards	11.2%	131.81	6.34
sports tickets	15.6%	1513.65	4.36
recreational equipment (music, sports)	2.8%	706.31	6.79
computing equipment	6.2%	789.51	6.09
multimedia	9.5%	541.26	6.46
jewelry	6.4%	409.09	6.36
other luxury	1.4%	199.46	4.3
automotive	18.4%	179.23	4.74
household items	6.1%	305.35	5.46
clothing and accessories	1.1%	166.12	6.33
Obs = 3000			
<i>cancelled</i>	<i>37.6%</i>	<i>489.07</i>	<i>5.85</i>

⁶⁶ Examination of the automotive parts category suggests that modification (hobby) was a likely motive since the retailer catered towards certain models of sports cars.

The most popular items were toys, consisting primarily of video game consoles, hand-held gaming devices and portable music players. Automotive parts were also a common layaway item. Sports tickets were also popular, consisting primarily of NFL season tickets. Multimedia equipment such as digital cameras, DVD or Blu-Ray players, and plasma televisions were another frequent category, as were gift cards to various retail establishments.

The fact that gift cards comprised one tenth of all layaway purchases is striking due to the fact that the arrangement is purely a shift of monetary resources over time, at a cost to the consumer. Although the layaway contractor did advertise a “promotion” for gift cards in July 2008, there was no discount associated with the promotion, and standard commissions and fees applied.⁶⁷ The fact that customers took up this promotion serves as supporting evidence that layaway has a self-control motivated appeal, rather than appeal due to misunderstanding of the rules of the layaway program.⁶⁸

Figure 3.2 below shows the distribution of number of purchases observed for each customer in the data set. Note that since we follow all customers who made their first purchase between June 2008 and June 2009, the chart only represents the accumulation of total purchases up until the data ends, although the customers may very well go on to make further contracts going forward.

⁶⁷ The idea for the promotion was inspired by the observation that many customers were already purchasing gift cards on layaway such that the gift cards would arrive in time for Black Friday.

⁶⁸ Note that a gift card is not in itself a consumption item – it must later be used to purchase an item which consequently yields utility, ruling out the idea consumers participated in layaway contracts only because they incorrectly thought that they could secure the availability of an item with a layaway contract. Also, since consumers could not see the future prices of items at the time that the gift cards would become available to them, it is not possible that they thought they could secure the price of an item using layaway.

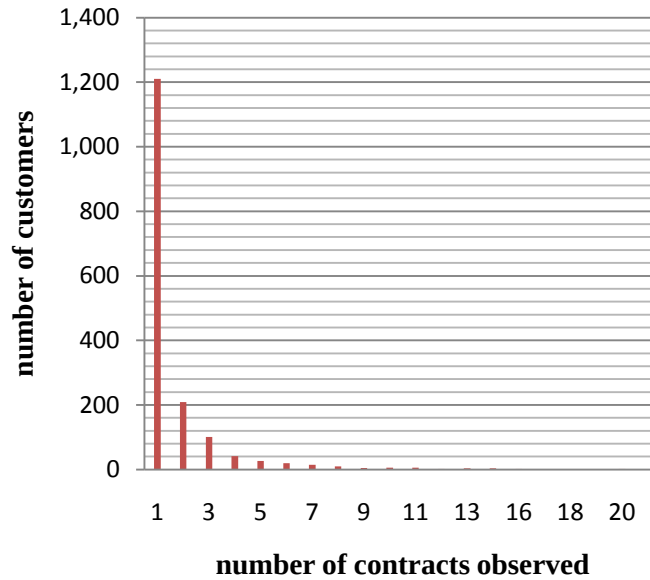


Figure 3.2: Distribution of number of purchases/contracts, by customer

3.5 Results

3.5.1 Contract choice

Table 3.2 below shows the full sample (temporarily ignoring any customer-specific effects) for the relationship between types of items ordered and chosen contract length. The comparison category is household items. The dependent variable is measured in months, which is the increment by which customers can select the desired time length of the contract. Dummy variables for month of expected arrival of the item are included in order to control for any differences in contract choice behavior by season.⁶⁹

The linear regression shows that contract length tends to increase with the price of the item, which supports the idea that customers purchased contracts for items which they could not afford to pay for in full. In other words, more expensive items generally

⁶⁹ For example, last minute holiday shopping may induce shorter contracts on items scheduled to arrive in those months.

required longer contract periods in order to complete the entire payment. “Total price” refers to the price of the item, including tax, shipping and layaway commission. Since these components are not a function of the contract length chosen, but only of the item's price, “total price” is simply a more complete measure of the item's cost.

Controlling for total price, the square of total price (to allow for non-linear price effects), toys, luxury items, sports tickets, and auto parts had significantly shorter contract lengths compared to household items, or in other words a higher per month commitment on the part of the consumer. Gift cards tended to have significantly longer contracts, requiring a lower per month commitment compared to household items.

Table 3.2: Contract length (dependent variable), full sample

	<i>Coefficient</i>	<i>standard error</i>	<i>t-statistic</i>	<i>p-value</i>
total price**	0.0010	0.0002	6.23	0.000
total price sq.**	0.0000	0.0000	-3.46	0.001
toys**	-0.9263	0.2441	-3.79	0.000
gift cards*	0.5675	0.2976	1.91	0.057
tickets**	-3.1114	0.2965	-10.49	0.000
other recreation	0.6345	0.4639	1.37	0.171
computing	-0.2971	0.3487	-0.85	0.394
other electronics	0.2425	0.2899	0.84	0.403
jewelry	0.2691	0.3465	0.78	0.438
other luxury**	-1.4741	0.4139	-3.56	0.000
auto parts**	-1.3262	0.2323	-5.71	0.000
clothing	0.3828	0.6766	0.57	0.572
constant**	5.7539	0.2188	26.30	0.000
comparison group: household items				
month of arrival dummy variables included (clustered std. errs.)				
Obs = 3000		Adj R-sq = 0.1298		
**significant at 95% level, *significant at 90% level				

The regression results in Table 3.3 show the relationship between item type and chosen contract length when including customer fixed-effects. The rationale for including fixed-effects is that customer heterogeneity can then be accounted for without detailed knowledge about a customer's income and other characteristics, identifying differences in

contract length across types of items based on the within-person variation. While this is beneficial since customers are in fact heterogeneous, the drawback is that only 458 customers in this particular sample made more than one contract over the sample period and the analysis will be reduced to those customers only.

The fixed-effects regression includes an estimation of outstanding payments for each consumer at the time of contracting to control for possible income effects on the contract length decision. The idea is that if a customer already has several layaway contracts in progress, being aware of his or her own budget constraints could leave the customer with little other choice but to choose a longer contract length.

The variable “outstanding payments” is an estimate and not necessarily the true amount, due to the fact that customers occasionally adjusted their payment schedule after signing up for a contract. For simplicity, and because I only observe the most recent payment schedule (and not any potential adjustments that occurred in between), I use the total price divided by the initial chosen contract length as the estimate for outstanding payments for months in which a payment was successfully made. In cases where the payment was not made for whatever reason, I assume that customers have rational expectations about future unfulfilled payments and therefore do not count those months as having an outstanding payment for that particular item.⁷⁰ Even though payment schedules can be customized to customer preference by specific day of the month (even multiple days per month) and not just on a fixed day of each month, I conduct all analysis on a monthly basis.

⁷⁰ Most of the unfulfilled payments were due to cancellations although occasionally there was an overdraft on the customer's bank account.

I also include a variable “contract ordering” which takes into account how many layaway contracts an individual customer has made thus far. Thus if the current observation is a customer's third layaway contract, the value of this variable is three. This variable is intended to capture any learning effects that may occur within-person.

The fixed-effects regression shows that within individual customers, toys, luxury items and auto parts still tended to induce shorter contract choice or relatively more commitment, while gift cards tended to induce longer contract choice hence relatively less commitment. Sporting event tickets, recreation items, computing goods, other electronics, and clothing were statistically indistinguishable from household items in the choice of contract length. Contract length was increasing in the price of the current item, and in the order of contracts implemented. Hence all else equal, learning appeared to be in the direction of less commitment per month, controlling for the price of the item and other layaway contracts in progress. The coefficient on outstanding payments is marginally negative, implying that month to month income constraints were not a concern on average.

The far right three columns show the statistical significance of the coefficients with standard errors clustered at the customer level. While having robust standard errors decreases the significance of the coefficients, it is not surprising since in most cases the coefficients are being identified on only two or three within-person observations.

3.5.2 Default and cancellation

Finally, we are interested in what features of the contract contributed to the likelihood of default or cancellation. To investigate this question, I limit the contract sample to those items which would have been fully paid for by December 2009, the date of the data retrieval. This leaves a total of 2501 contracts in the probit regression. Table 3.4 shows the probability of cancellation of the order based on the variables listed.

I am able to observe whether the order was fulfilled or not, but not the exact reason for the order not being fulfilled. For example, I cannot tell whether the customer canceled the order due to an increase in price of the item, whether he or she changed their mind about wanting the item, or whether the order was canceled due to repeated insufficient funds in the customer's bank account.

As expected, longer contracts and items of higher price were more likely to be canceled. The likelihood of cancellation was decreasing in the amount of outstanding payments, and increasing in the number of contracts initiated. Gift cards, sports tickets and auto parts were significantly less likely to be canceled compared to household goods. Computing equipment was more likely to be canceled compared to household goods. The cancellation results may reflect the resale potential of the various items, but a self-control or household bargaining story is not immediately obvious.

Table 3.4: Likelihood of cancellation/default

	<i>Coefficient</i>	<i>standard error</i>	<i>z-score</i>	<i>p-value</i>
contract length**	0.0430	0.0127	3.37	0.001
contract ordering**	0.0646	0.0113	5.73	0.000
outstanding payments**	-0.0028	0.0003	-8.34	0.000
total price**	0.0003	0.0001	3.47	0.001

Table 3.4: Likelihood of cancellation/default, Continued

total price sq.*	0.0000	0.0000	-1.74	0.082
toys	-0.1627	0.1067	-1.52	0.127
gift cards**	-0.4477	0.1266	-3.54	0.000
tickets**	-1.6006	0.1504	-10.64	0.000
other recreation	0.0751	0.1824	0.41	0.680
computing**	0.4384	0.1530	2.87	0.004
other electronics	0.1316	0.1254	1.05	0.294
jewelry	-0.1476	0.1433	-1.03	0.303
other luxury	0.2913	0.2301	1.27	0.205
auto parts**	-0.2918	0.1139	-2.56	0.010
clothing	0.1358	0.2647	0.51	0.608
January	-0.0153	0.1585	-0.10	0.923
February*	0.2897	0.1627	1.78	0.075
March	0.1736	0.1544	1.12	0.261
April*	0.2816	0.1434	1.96	0.050
May**	0.5327	0.1176	4.53	0.000
July**	0.2453	0.1198	2.05	0.041
August**	0.4201	0.1320	3.18	0.001
September*	0.2560	0.1413	1.81	0.070
October	0.0950	0.1441	0.66	0.509
November**	0.4320	0.1272	3.40	0.001
December**	0.3288	0.1308	2.51	0.012
constant**	-0.6443	0.1450	-4.44	0.000
comparison group: household items				
comparison month: June				
Obs = 2501, Log likelihood = -1400.02				
**significant at 95% level, *significant at 90% level				

3.6 Conclusions

This paper analyzed a data set of layaway plans with flexible contract lengths, finding that compared to household items, items that were entertainment-oriented such as video games and multimedia devices, as well as luxury goods, tended to induce shorter contract lengths. Controlling for price, this corresponded to a higher month to month commitment to not spend on very short-term indulgences by having a higher dollar amount automatically withdrawn from the customer's account each month. The result was

robust across individual customers, remaining marginally significant within individual repeat customers.

A natural explanation is that people are simply eager and willing to save for items which are more enjoyable for them. They are more reluctant and less enthusiastic about saving for mundane items such as household appliances or clothing.⁷¹ This is consistent with the longer contract lengths observed for gift cards, since gift cards do not provide a specific savings target that the customer can envision and anticipate enjoying.

In other words, people appear to be sophisticated in this setting about saving, but seemed to use that sophistication more aggressively towards fulfilling relative indulgences. If this type of commitment behavior is common in other field settings, it raises a broader question of whether free form commitment devices will be used as a policymaker might have intended or whether consumers may possibly find innovative ways to use them while defeating the original self-control motive.⁷²

The contribution of this study to the savings literature is that it examines a savings commitment environment in which the item being saved for is explicitly observable. This information sheds light on consumers' relative priorities for various types of items in short term savings devices. Further work in this area might examine medium-term savings priorities and success rates among households.

⁷¹ Another potential explanatory variable (which I do not observe in this particular dataset) is price and preference uncertainty. Having uncertainty about future prices or personal preferences for an item should generally reduce the desire to commit. Hence from a monthly sacrifice perspective, consumers may be more reluctant to commit a large sum, and thus want to stretch the payments over a longer time horizon (this assumes that the distribution on prices and preferences is not time-varying). Although price changes do not seem to be too prevalent in this data, the preference uncertainty interpretation of the results would suggest that consumers were more certain about wanting entertainment and luxury items in the future compared to household items.

⁷² For example, participants in a smoking cessation program with a nicotine urine test may turn to other harmful substances to help curb their desire for cigarettes.

Chapter 3 is being prepared for submission for publication. The dissertation author was the primary investigator and author of this material.

Appendix 3.A: Categorization of items

Toys

- MP3/MP4 Players (iPod Nano, iPod Touch, iPod Shuffle, Zune, others)
- Microsoft Xbox 360 console and accessories
- Nintendo DS, Wii and accessories/games
- Sony Playstation 3/PSP and accessories/games
- other miscellaneous toys

Gift cards

- most popular: Wal-Mart, Target, Old Navy, Best Buy
- several others (AMC theaters, Amazon.com, JC Penney, Lowes, Starbucks, Williams-Sonoma, Macy's, etc)

Sports Tickets

- NFL season tickets

Other recreation

- hunting equipment
- fishing equipment
- musical instruments
- paintball equipment
- hotel rooms

Computing

- netbooks and laptops
- desktop computers
- memory cards
- software

Multimedia

- digital cameras
- LCD/HD televisions
- GPS
- cell phones
- camcorders

Jewelry

Other luxury

- designer handbags
- perfume

Household items

- bedding
- refrigerators
- cooking devices
- items for baby
- clocks
- air purifiers
- tool sets

Automotive parts

Clothing and accessories

- clothing
- watches
- bags/ backpacks
- wallets
- shoes
- sunglasses

Appendix 3.B: Commission effects

Table 3.B.2: Contract length (dependent variable), full sample

	<u>Coefficient</u>	<u>standard error</u>	<u>t-statistic</u>	<u>p-value</u>
total price**	0.0010	0.0002	6.23	0.000
total price sq.**	0.0000	0.0000	-3.48	0.000
commission indic	0.2679	0.1875	1.43	0.153
toys**	-0.9327	0.2441	-3.82	0.000
gift cards*	0.5771	0.2976	1.94	0.053
tickets**	-3.1388	0.2964	-10.59	0.000
other recreation	0.6239	0.4630	1.35	0.178
computing	-0.3105	0.3485	-0.89	0.373
other electronics	0.2487	0.2898	0.86	0.391
jewelry	0.2578	0.3472	0.74	0.458
other luxury**	-1.4515	0.4186	-3.47	0.001
auto parts**	-1.3549	0.2331	-5.81	0.000
clothing	0.3664	0.6735	0.54	0.586
constant**	5.5265	0.2734	20.21	0.000
comparison group: household items				
month of arrival dummy variables included (clustered std. errs.)				
Obs = 3000 Adj R-sq = 0.1302				
**significant at 95% level, *significant at 90% level				

Table 3.B.4: Likelihood of cancellation/default

	<u>Coefficient</u>	<u>standard error</u>	<u>z-score</u>	<u>p-value</u>
contract length**	0.0441	0.0128	3.45	0.001
contract ordering**	0.0630	0.0113	5.56	0.000
outstanding payments**	-0.0028	0.0003	-8.30	0.000
total price**	0.0003	0.0001	3.43	0.001
total price sq.*	0.0000	0.0000	-1.70	0.089
commission indic**	0.2768	0.0904	3.06	0.002
toys	-0.1711	0.1071	-1.60	0.110
gift cards**	-0.4356	0.1271	-3.43	0.001
tickets**	-1.6372	0.1511	-10.84	0.000
other recreation	0.0639	0.1823	0.35	0.726
computing**	0.4213	0.1534	2.75	0.006
other electronics	0.1356	0.1257	1.08	0.281
jewelry	-0.1639	0.1434	-1.14	0.253
other luxury	0.3029	0.2310	1.31	0.190
auto parts**	-0.3262	0.1146	-2.85	0.004
clothing	0.1241	0.2647	0.47	0.639
January	-0.0205	0.1584	-0.13	0.897
February*	0.2871	0.1627	1.76	0.078
March	0.1699	0.1545	1.10	0.271
April**	0.2815	0.1434	1.96	0.050
May**	0.5266	0.1174	4.48	0.000
July**	0.2414	0.1197	2.02	0.044
August**	0.4112	0.1322	3.11	0.002
September*	0.2437	0.1417	1.72	0.085
October	0.0857	0.1439	0.60	0.551
November**	0.4360	0.1273	3.43	0.001
December**	0.3201	0.1308	2.45	0.014
constant**	-0.8726	0.1633	-5.34	0.000

comparison group: household items

comparison month: June

Obs = 2501, Log likelihood = -1395.28

**significant at 95% level, *significant at 90% level

References:

- Anderson, S., and Baland, J., “The Economics of Roscas and Intrahousehold Resource Allocation”, *The Quarterly Journal of Economics*, Vol. 117, No. 3 (Aug 2002), pp. 963 – 995.
- Ashraf, N., Karlan, D., Yin, W., “Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines”, *The Quarterly Journal of Economics*, Vol. 121, No.2 (May 2006), pp. 635 – 672.
- Bryan, G., Karlan, D., and Nelson, S. “Commitment Devices”, Working Paper, April 2010.
- DellaVigna, S. and Malmendier, U., “Paying Not to Go to the Gym”, *American Economic Review*, Vol. 96, No. 3 (2006) pp. 694 – 719.
- Fudenberg, D. and Levine, D.K., “A Dual-Self Model of Impulse Control”, *American Economic Review*, Vol. 96, No. 5 (2006) pp. 1449-1476.
- Gine, X., Karlan, D.S., and Zinman, J., “Put Your Money Where Your Butt is: A Commitment Contract for Smoking Cessation”, World Bank Policy Research Working Paper No. 4985. July 2009.
- Gul, F. and Pesendorfer, W., “Temptation and Self-Control”, *Econometrica*, Vol. 69, No. 6 (Nov 2001), pp. 1403 – 1435.
- Laibson, D., “Golden Eggs and Hyperbolic Discounting”, *The Quarterly Journal of Economics*, Vol. 112, No. 2, May 1997, pp 443-477.
- Oster, S.M., and Scott Morton, F.M. (2005) “Behavioral Biases Meet the Market: The Case of Magazine Subscription Prices” *Advances in Economic Analysis and Policy*: Vol. 5: Iss. 1, Article 1.
- Strotz, R.H., “Myopia and Inconsistency in Dynamic Utility Maximization”, *The Review of Economic Studies*, Vol. 23, No. 3 (1955-1956), pp.165-180.
- Thaler, R.H., and Shefrin, H.M., “An Economic Theory of Self-Control”, *The Journal of Political Economy*, Vol.89, No. 2 (Apr. 1981), pp. 392-406.