Choice Overload and Simplicity Seeking

Sheena S. Iyengar
Columbia University, Graduate School of Business

Emir Kamenica
University of Chicago, Graduate School of Business

February 6, 2007

ABSTRACT

In settings such as investing for retirement or choosing a drug plan, individuals face a staggering number of options. In this paper, we analyze how an abundance of options influences which alternative is selected. We present both laboratory experiments and field data that confirm our theoretical prediction: larger choice sets induce a stronger preference for simple, easy-to-understand options. The first experiment demonstrates that, in seeming violation of the weak axiom of revealed preference, subjects are more likely to select a given sure bet over non-degenerate gambles when choosing from a set of 11 options than when choosing from a subset of 3. The second experiment clarifies that excessive choice sets induce a preference for simpler, rather than less risky, options. Lastly, using records of more than 500,000 employees from 638 institutions, we demonstrate that the presence of more funds in an individual's 401(k) plan is associated with a greater allocation to money market and bond funds at the expense of equity funds.

1. Introduction

A growing body of research in psychology and economics demonstrates that agents can be better off with a strictly smaller choice set. Evidence from both the laboratory and the field indicates that a person's willingness to participate in a market, e.g., purchase a good (Iyengar and Lepper 2000, Boatwright and Nunes 2001), take up a loan (Bertrand *et al.* 2005), or enroll in a 401(k) plan (Iyengar, Huberman, and Jiang 2004), can decrease when participation requires selecting an alternative from a larger set of options. Theoretical investigations show that this reversal of a simple choice-theoretic principle that "more is better" may arise when the availability or absence of an option conveys useful information (Kamenica 2006, Kuksov and Villas-Boas 2005) or when agents exhibit preferences with regret (Irons and Hepburn 2003, Sarver 2005). This previous research on choice overload has focused exclusively on how increasing assortment size affects *whether* an agent will participate in a market. In contrast, the present paper investigates how increasing assortment size affects *which* alternative an agent will choose.

This question is of particular interest in the context of decisions such as investing for retirement or selecting a drug plan since: (i) those decisions are of great consequence, (ii) the choice sets are typically extensive, (iii) the majority *does* save for retirement and participate in a drug plan, despite the fact some get discouraged by the abundance of alternatives, and (iv) government policy can have an enormous impact on the number of options that individuals face.

Contextual inference theory (Kamenica 2006) suggests that an individual presented with a larger number of options will be more likely to select simpler

alternatives, i.e., those alternatives whose utility is more transparent to the decision maker. The basic intuition for this result is straightforward: under plausible regularity conditions, larger product lines yield lower average utility in a market equilibrium. Hence, when faced with large choice sets, agents are reluctant to select options that they do not understand well. As a result, excessive choice sets induce a preference for simplicity.

We examine this prediction both in the laboratory and in the field. In the first experiment, we offer subjects a choice of either 11 or 3 gambles. Specifically, one group of subjects selects a lottery from the menu of 11 gambles which includes 10 risky options and one degenerate lottery. Another group of subjects is offered a 3-gamble subset of the 11 gambles which includes the degenerate lottery. As predicted by contextual inference theory, regularity is violated and more subjects choose the simple option ($5 for sure) when presented with the set of 11 options than when presented with the set of 3.

In the first experiment, the simplest option is also the least risky. The theory suggests, however, that larger choice sets induce a greater preference for *simpler* options, not for less risky ones. The second experiment removes the confound of riskiness and complexity by using a design where the simplest option is the riskiest. In particular, once again subjects are offered either 11 gambles or a subset of 3. The set of 11 includes 10 gambles which yield a distinct amount of money (between $0 and $10) for each outcome of a die toss (e.g., $1.50 if ⚀, $9.25 if ⚁, $8.75 if ⚂, $7.00 if ⚃, $0.75 if ⚄, $1.25 if ⚅). The remaining gamble yields $0.00 if the die indicates ⚀, ⚁, or ⚂, and $10.00 if it indicates ⚃, ⚄, or ⚅. This all-or-nothing gamble is both simpler and riskier than the

other 10. As predicted, the simpler gamble is selected more often from the larger choice set than from the smaller choice sets.

Finally, using data from the Vanguard Center for Retirement Research, we analyze the investment decisions of over 500,000 employees in 638 firms**.** We use the variation in the number of funds across 401(k) plans to investigate how employees respond to the number of investment options they face. Conditional on a host of individual and plan-level controls, with every additional 10 funds in a plan, allocation to equity funds decreases by 3.28 percentage points. Moreover, for every 10 additional funds, there is a 2.87 percentage point increase in the probability that participants will allocate nothing at all to equity funds. We address the concerns about potential endogeneity and selection biases by demonstrating that those observable characteristics that predict a greater number of funds are typically associated with a greater exposure to equities. Accordingly, controlling for observable individual and plan-level characteristics strengthens our results.

The next section provides a broader discussion of choice overload. Section 3 discusses the findings from the laboratory experiments. Section 4 describes the 401(k) data and analyzes how the number of funds affects asset allocation. Section 5 concludes.

2. Choice Overload: Theory and Evidence

A simple choice-theoretic principle requires that, in non-strategic situations, an agent always obtains weakly greater welfare from a larger set of options.[1] However, several recent papers show that consumers sometime prefer strictly smaller choice sets

[1] Exceptions to this principle on grounds of temptation and self-control problems (e.g., Laibson 1997, Gul and Pesendorfer 2001) have spawned a productive line of research, but we ignore these issues for our purposes.

(Iyengar and Lepper 2000, Boatwright and Nunes 2001, Bertrand *et al.* 2005). A study by Iyengar, Huberman, and Jiang (2004) reports that, controlling for individual- and plan-level attributes, increasing the number of funds available in a 401(k) plan decreases the percentage of employees who participate in the plan.[2] Employees in plans that offer 5 funds have a predicted participation probability of 72%. When the number of funds in the plan increases to 35 however, the predicted participation probability drops to 67.5%. With 56 funds, participation drops to 61%. On average, for every additional 10 funds available, the predicted individual participation probability declines by about 2 percentage points.

Why might an individual prefer a smaller choice set? One reason might be the possibility of regret. As Irons and Hepburn (2003) and Sarver (2005) argue, if agents experience sufficient regret upon finding out that the option they selected was not *ex post* optimal given the choices they had available, they might prefer to limit their choice set so as to eliminate the discomfort of regret. This explanation seems of limited relevance, however, since choice overload has typically been demonstrated in settings where consumers never find out how *ex post* successful their initial choice had been. Moreover, Salgado (2005) explicitly tests the importance of regret for choice overload and reports no impact on the preference for smaller choice sets.

Another explanation of choice overload is through contextual inference theory (Kamenica 2006). Consider a situation where some people are unsure about the utility of some options. Call the alternatives whose utility is transparent to all *simple options* and

[2] Huberman, Iyengar, and Jiang (2004) provide a more detailed discussion of the broader determinants of participation and contribution rates in 401(k) plans.

the alternatives whose utility some find hard to gauge *complex options*.[3] To be precise, suppose that the utility a person i receives from option x is $u_i(x,s_x)$. Then, if s_x is common knowledge, we say x is simple. On the other hand, if some people do not know s_x we say x is complex.[4]

A common feature of choice sets that arise as product market equilibria is that they contain precisely those goods that yield the greatest average utility. As Kamenica (2006) demonstrates, if markups are constant and valuations for alternatives are uncorrelated, the unique equilibrium product line exhibits this property. An immediate and important consequence of this feature of equilibrium choice sets is that random selection from a larger menu yields lower expected utility than random selection from a smaller menu. Selecting an option of uncertain utility is fine when only a few options are available since all elements of a small choice set yield high expected utility. In contrast, choosing an alternative you are uncertain about may be unwise when there are many options, since the large choice set also includes niche products that yield low average utility. To avoid the possibility of selecting such a product, it is better to select a simple option when choosing from a large choice set.[5]

The analysis in this paper draws on the notion that people act as if the choice sets they encounter across a variety of decision contexts share this features of equilibrium product lines. Levitt and List (2006) emphasize the importance of the possibility that

---

[3] We acknowledge that it would be more useful to have a definition of complexity based on the attributes of the options, but providing an attribute-based definition applicable across domains would require its own research agenda.

[4] Alternatively, we could say x is complex if obtaining information about s_x is costly. Also, one could easily define an analogous continuous definition of complexity.

[5] Note that cognitive costs or search costs *per se* make no such prediction. If small choice sets are a random selection from large choice sets, the relative attractiveness of simple options is independent of the number of alternatives. Contextual inference theory provides a microfounded basis for the difference between small choice sets and random selections from large choice sets.

experimental subjects bring context from everyday life into the lab. In the decision problems we examine in this paper, such transfer of context from the market setting to the lab or to the 401(k) allocation decision would lead a person to be more likely to select a simple, easy-to-understand option when presented with a greater number of alternatives. The remainder of this paper tests whether people indeed do so.

3. The Experimental Tests

The experiments were conducted at Columbia University. Research assistants randomly approached passers-by on or near the university campus and requested their participation in completing a brief one-page questionnaire, the content of which was unrelated to the experimental manipulations. Of the people approached, 90% agreed to participate.

After completing the survey, the subjects were offered a set of gambles, from which they chose one as compensation. The subjects assigned to the Extensive condition were offered a set of 11 gambles, while those in the Limited condition were offered a subset of 3 gambles out of the original 11. The order in which the gambles were presented was randomized across subjects and within research assistants. The subjects were not aware of the existence of the other condition. The gambles were constructed so that those with higher expected values have a higher variance,[6] and had similar, though not identical, prospect theoretic values as calibrated using estimates from Kahneman and Tversky (1993). The two experiments only differed in the structure of the gambles. The

[6] This relationship is not strict, however. Due to rounding issues (we did not want to pay subjects in denominations less than 25 cents), there are pairs of gambles with the same expected value and slightly different levels of risk.

relevant excerpts from the instructions (excluding the unrelated questionnaire) are provided in the Appendix.

Experiment 1.

In the first experiment, subjects were provided with a menu of binary gambles. After the subject selected the gamble, the experimenter would flip a coin to determine the amount of the subject's compensation. Each gamble was described by the dollar amount the subject would receive should the coin fall heads and the amount should it fall tails. Of the 11 gambles, 10 were non-degenerate (e.g., if the coin indicates 'heads,' the subject gets $4.50; if the coin indicates 'tails,' (s)he gets $7.75) and one was a sure bet ($5.00 whether the coin falls on 'heads'or 'tails'). The 3 gambles presented to the subjects in the Limited condition included the degenerate lottery.[7] Table I provides the lists of gambles in the two conditions. Note that, as the instructions included in the Appendix reveal, the simple gamble was presented in the same format as the other gambles and embedded in the list of the other gambles. Hence, in neither condition did the simple option stand out in the presentation of the gambles.

Results. We observe a strong violation of regularity. Only 16% of the 69 subjects in the Limited condition chose the $5 for sure, but 63% of the 68 subjects in the Extensive condition did so (Fisher's exact p-value < 0.001). Figure 1 depicts the histogram of the distribution of choices in the two conditions. The magnitude of this difference is striking: subjects were roughly four times more likely to select the $5 for sure when facing 10 other options than when facing only 2.

[7] The other two gambles were selected randomly but were constant across subjects. Those two additional gambles were not atypically attractive, based on the choices in the Extensive condition. In Experiment 2, we remedy this issue and randomly select the two additional gambles independently for each subject.

Experiment 2.

In the first experiment, the simplest option was also the least risky. Experiment 2 is designed to confirm our prediction that it is the simplicity, rather than the lower risk, that becomes more attractive as the size of the choice set grows. Hence, this time there were six possible outcomes associated with each gamble, and compensation was determined by a die toss. Of the 11 gambles, 10 yielded a distinct amount[8] (between $0 and $10) for each outcome of the die toss (e.g., if the die falls on ⚀, the subject receives $4.25; if ⚁, $5.50; if ⚂, $9.75; if ⚃, $8.50; if ⚄, $0.00; if ⚅, $0.75), while one was riskier, paying out either $0 (on a ⚀, ⚁, or ⚂) or $10 (on a ⚃ , ⚄, or ⚅). The subset of 3 gambles presented to the subjects in the Limited condition always included the simpler, all-or-nothing bet, as well as two other gambles which were independently and randomly selected for each subject from the other ten gambles in the Extensive condition. As in the previous experiment, the simpler gamble did not stand out in the presentation. Table II provides the lists of gambles in the Extensive condition.

Results. Once more, the simpler gamble was selected more frequently from a larger choice set. Only 16% of the 62 subjects in the Limited condition chose the simplest gamble, while 57% of the 58 Extensive condition subjects did so (Fisher's exact p-value < 0.001). Figure 2 illustrates the histogram of the distribution of choices in the two conditions.

Jointly, these experimental findings support the prediction of contextual inference theory that increasing the size of the choice set strengthens the appeal of easier-to-understand options.

[8] The amount for each outcome was selected *with* replacement, so even in these 10 gambles there is sometime the same payout for two distinct outcomes of the die toss.

4. Fund Availability and Asset Allocation

In this section, we examine whether the mechanism underlying the experimental results affects behavior in a real world setting with large stakes. In particular, we examine employees' asset allocation decisions for their 401(k) plans.

Defined contribution plans, such as 401(k) plans, are one of the primary instruments for retirement savings. A 401(k) plan is an institutional plan in which employees are given monetary incentives to transfer some of their salary into investment funds provided by the plan. Participating in a 401(k) plan affords employees tax deferred income and employer match, i.e., employers will often match at least some of their employees' retirement plan contributions. Over the years, a trend towards replacing pension plans with defined contribution plans has been observed, with a three-fold increase in 401(k) plan account assets between 1990 and 2001. In 2001, 45 million participating American workers held a total of $1.75 trillion in their 401(k) plans (Holden and Van Derhei 2003).

Upon joining a 401(k) plan, an individual must choose both how much income to contribute to the plan and how to allocate the investment across the funds the plan provides. In our data, the number of funds available in a plan varies from 4 to 59.[9] This variation allows us to examine how the number of options affects asset allocation.

4.1 The Data

The data was provided by the Vanguard Center for Retirement Research, whose records from 2001 include 639 defined contribution (DC) pension plans[10] with 588,926

[9] In the raw data, for one single observation, the number of funds was apparently miscoded as 2. Under the inclusion criteria specified below, we exclude that observation from our analysis.
[10] For the purpose of this analysis, a plan refers to an institution which provides 401(k) plans to eligible employees.

participating employees. Most of the DC plans offered Vanguard fund options and many offered funds from other fund families as well, e.g., TIAA/CREF. The data includes all the individuals regardless of how they chose to invest their money. An important feature of this choice making environment was that employees could not turn to plan providers for explicit advice on which funds to invest in. In fact, existing 401(k) education materials purposefully avoid recommending specific plans so as to escape ERISA classification as investment advice (Mottola and Utkus 2003).[11] Those individuals who choose to participate specified the deferral rate at which they would contribute (i.e., the percentage of their pre-tax annual salary they wish to contribute) and were provided with a list of funds to which they could allocate their investment.

At the individual level, the data provides information pertaining to gender, age, tenure, compensation, and wealth. Based on this information, we define self-explanatory variables *Female, Age, Age*^{2}*, Tenure, Tenure*^{2}*, logCompensation, and logWealth.*[12] Employees' wealth was measured through their IXI index (ranging from 1 to 24), based on the employees' 9-digit Zip Codes.[13]

We exclude individuals whose income is below $10,000 or above $1,000,000 as well as those below 18 years of age. We also exclude 165 individuals who did not contribute a positive amount to their 401(k) plan or had withdrawn money from any fund type. These criteria leave us with 580,855 employees in 638 distinct funds.

[11] The Employee Retirement Income Security Act of 1974 (ERISA) is a federal law that created minimum standards by which most voluntarily established pension and health plans in private industry must provide protection for individuals in these plans.

[12] For the observations where gender was missing, we use the percentage of female employees in the plan. The results are unchanged if we instead drop those observations.

[13] A company called IXI collects retail and IRA asset data from most of the large financial services companies, receiving data from all companies at the 9-digit Zip Code. IXI then assigns a wealth rank (from 1 to 24) to the Zip Codes, based on imputed average household assets.

Our dependent variables are measures of the employees' contributions across different types of funds. Our data provides information on how individuals allocate their total annual 401(k) contribution in 2001 (including both employee and employer match) into seven different categories: money market funds, bond funds, balanced funds, active stock funds, indexed stock funds, company stock, and other funds (mainly insurance policies and non-marketable securities). Variable *Equity%* indicates what fraction of the total 2001 contribution an employee placed in all types of equity (active funds, balanced funds, company stock, and index funds). We exclude "other" funds from the analysis, but since only 96 employees allocated a positive amount to these funds, any categorization of "other" leaves results unchanged. Finally, while we arbitrarily categorize balanced funds as equity, any other categorization (e.g., counting them as ½ bond and ½ equity, or counting them fully as bond funds) weakly strengthens our results. We also define an indicator variable *NoEquity,* which takes the value of 1 if an employee allocates her entire contribution only to money market and bond funds, and value 0 otherwise. Table III reports the summary statistics at the individual level.

At the plan level the data provides information about specific attributes of the retirement savings program. Employers in 538 out of 638 plans (covering 89% of the employees in the sample) offered some match to their employees' contributions. The match rates ranged from 0% to 250%, with most falling between 50% and 100%. Variable *Match* indicates the employer's match rate. In our sample, 102 plans (covering 53% of the sample employees) had "own-company stock" as an investment option. Variable *CompanyStockOffered* is an indicator variable for whether the plan offers company stock, and *RestrictedMatch* is an interaction variable equal to zero if the match

is with company stock only and equal to *Match* otherwise. The Defined Benefits plan is a company pension plan in which retired employees receive specific amounts based on salary history and years of service while their employers bear the investment risk. Contributions to a DB plan may be made by the employee, the employer, or both. In our data, 215 plans (covering 62% of the employees in this sample) offered defined-benefit options. Indicator variable *DBPlanOffered* indicates whether a defined benefits plan was available. Almost all of the plans offered Web accessibility. Variable *PercentWebUse* represents the percent of plan participants who registered for web access to their 401(k) accounts.[14] Variable *logNumberEmployees* captures the size of the firm.

We also define plan level variables that capture aggregate measures of the attributes of employees in the plan. The self-explanatory variables *logPlanAverageCompensation*, *logPlanAverageWealth*, *PercentFemale*, *PlanAverageAge*, and *PlanAverageTenure* allow us to provide some control for plan-level policies that depend on the aggregate characteristics of people within the plan.

Finally, our key independent variable of interest is *NumberOfFunds*, the number of funds offered by a plan. The median number of funds per plan was 11. Ninety percent of plans offered between 6 and 25 fund choices, and 18 plans offered 30 options or more. Table IV provides summary statistics at the plan level.

4.2 The Impact of Number of Funds on Allocation

We first consider OLS regression:

(1) $$Equity\%_{ij} = \beta_0 + \beta_1 * NumberOfFunds + \beta_2 * X_{ij} + \beta_3 * Z_j + \varepsilon_{ij},$$

[14] As Choi, Laibson, and Metrick (2002) observed, web access may have a significant impact on trading frequency for 401(k) participants.

where $Equity\%_{ij}$ is the percentage of contribution individual *i* in plan *j* allocated to equities, X_{ij} is the vector of individual-specific attributes:{*Female, Age, Age*^{2}*, Tenure, Tenure*^{2}*, logCompensation, logWealth* }, and Z_j = {*Match, CompanyStockOffered, RestrictedMatch , DBPlanOffered, PercentWebUse, logNumberEmployees, logPlanAverageCompensation, logPlanAverageWealth, PercentFemale, PlanAverageAge, PlanAverageTenure*} denotes plan-level characteristics.

Column (1) of Table V reports the impact of the number of funds on the percentage of contribution to equities. On average, for every 10 funds added to a plan, the allocation to money market and bond funds increases by 3.28 percentage points (t=2.81), at the expense of equities. Given the mean allocation to money market and bond funds of 22.16%, this effect is of limited, but not negligible, economic significance.

Column (2) examines the impact of the size of the choice set on the probability that an employee invests no money whatsoever in equity funds, using a linear probability model:

(2) $$NoEquity_{ij} = \beta_0 + \beta_1 * NumberOfFunds + \beta_2 * X_{ij} + \beta_3 * Z_j + \varepsilon_{ij} .$$

While on average only 10.53% of employees do not invest any money in equities, this probability increases by 2.87 percentage points for every 10 additional funds (t=2.76).[15]

These results are all the more striking because the percentage of funds that are equity funds increases in the overall number of funds.[16] In fact, both the fraction of

[15] For both specifications (1) and (2), we have also examined whether the effect of the number of funds on allocation is weaker for more sophisticated employees (proxied by age, tenure, income, or wealth) but find no evidence for heterogeneous effects.

[16] For the median plan, ¾ of the funds are equity funds, and this percentage increases by 3.94 percentage points for every 10 additional funds.

contributions allocated to equity funds and the probability that at least some funds are allocated to equities are *decreasing* in the number of equity funds.[17]

4.3 Endogeneity and Selection

The interpretation of the results in Table V is complicated by both the possibility that plans with different demographics have different number of funds, and the possibility that the type of employee that self-selects into participation varies with the number of funds in a plan. To address these two concerns, we compare the manner in which *Equity%* and *NoEquity* vary with the individual- and plan-level attributes with the way in which *NumberOfFunds* does so. Specifically, we compare the coefficients β_1 and β_2 across these OLS regressions:

(3) $$Equity\%_{ij} = \beta_0 + \beta_1 * X_{ij} + \beta_2 * Z_j + \varepsilon_{ij},$$

(4) $$NoEquity_{ij} = \beta_0 + \beta_1 * X_{ij} + \beta_2 * Z_j + \varepsilon_{ij},$$ and

(5) $$NumberOfFunds_{ij} = \beta_0 + \beta_1 * X_{ij} + \beta_2 * Z_j + \varepsilon_{ij}.$$

Table VI reports the results. Unsurprisingly, equity exposure is generally more correlated with individual characteristics, while the number of funds varies more closely with plan-level attributes. Moreover, the only covariates that have a significant impact on both allocation to equity and the number funds affect the two in the same direction. Consequently, the inclusion of controls either has no effect on the coefficients on *NumberOfFunds* in regressions (1) and (2) or strengthens our results. Of course, without exogenous variation in the number of funds, we cannot with certainty exclude the

[17] Huberman and Jiang (2006) find a positive, though small and only marginally significant, relationship between exposure to equities and the *fraction* of equity funds. Their analysis is most supportive of Benartzi and Thaler's (2001) 1/n hypothesis when there are fewer than 10 funds offered.

possibility that some unobservable covariates drive the relationship in Table V, but the results in Table VI greatly alleviate this concern. To the extent that observable characteristics are representative of the unobservables, the omitted variable bias is likely to weaken our results. In particular, the coefficients in Table VI suggest that unobserved employee sophistication would be associated with more funds and greater equity exposure.

In principle, one could further address the issue of endogeneity by instrumenting for the number of funds in an employee's plan with the individual characteristics of other employees in the firm,[18] but the relationship between aggregate characteristics of the plan and its number of funds is too weak.[19] Overall, the observed patterns in asset allocation, together with the experimental findings, support the view that larger choice sets induce a stronger preference for simpler options.

5. Conclusion

Previous research on choice overload focuses exclusively on the possibility that increasing the number of options can reduce participation in a market. This investigation moves beyond prior studies by examining the way in which the informational content of choice sets impacts which alternatives are selected from large assortments.

As Thaler and Benartzi (2004) have demonstrated, decisions about retirement savings can be highly influenced by the structure of the savings program. Our results indicate that the number of funds offered is an additional feature of savings plans that can

[18] The possibility of social influence would compromise the validity of this instrument, though these peer effects would likely bias the estimates toward zero. See Duflo and Saez (2003) for evidence on social influence in retirement plan participation.
[19] Importantly, the results of this two-stage least squares estimation cannot reject the estimates in Table V.

have considerable impact on behavior. Given the prevalence of large choice sets and the sensitivity of the size of choice sets on government policy (e.g., Social Security reform, Medicare reform), the potential implications of our results are substantial. Many politicians in the United States are advocating a partial privatization of the Social Security system. Such a change would dramatically increase the number of investment options for retirement that individuals would have to choose from. This increase, unless accompanied with an effort to provide suitable information, could substantially impact the structure of pension portfolios, and in turn, the expected resources available to future retirees.

The question of how to provide useful information that would allow employees and consumers to benefit, rather than suffer, from the abundance of options that modern knowledge and technology provide thus becomes of great importance. Future research that takes into account complicating issues that are beyond the scope of this paper, such as regulatory capture and potential inefficiencies of regulation, would allow us to construct specific policy suggestions.

Table I: Set of Gambles for Experiment 1

Extensive Condition		
Gamble #	If heads	If tails
1	$5.00	$5.00
2	$4.50	$7.75
3	$4.00	$8.25
4	$3.50	$8.75
5	$3.00	$9.50
6	$2.50	$10.00
7	$2.00	$10.50
8	$1.50	$11.25
9	$1.00	$11.75
10	$0.50	$12.50
11	$0.00	$13.50

Limited Condition		
Gamble #	If heads	If tails
1	$5.00	$5.00
2	$3.50	$8.75
3	$0.00	$13.50

Table II: Set of Gambles for Experiment 2

Extensive Condition						
Gamble #	If ⚀	If ⚁	If ⚂	If ⚃	If ⚄	If ⚅
1	$0.00	$0.00	$0.00	$10.00	$10.00	$10.00
2	$1.50	$9.25	$8.75	$7.00	$0.75	$1.25
3	$4.25	$5.50	$9.75	$8.50	$0.00	$0.75
4	$1.00	$2.00	$6.75	$7.50	$5.75	$4.75
5	$5.50	$1.00	$0.75	$6.50	$7.50	$6.75
6	$0.00	$0.00	$8.75	$2.75	$9.75	$8.00
7	$9.75	$3.00	$7.00	$6.50	$0.50	$1.50
8	$9.50	$1.50	$1.50	$2.50	$3.25	$10.00
9	$5.50	$8.50	$3.25	$0.00	$8.50	$2.50
10	$9.25	$7.75	$3.75	$2.00	$3.25	$2.00
11	$1.25	$4.50	$8.50	$8.75	$4.50	$0.75

Table III: Descriptive Statistics of the Employees

	Mean	St. Dev.	Min	Max	Obs.
Female	0.38	0.46	0	1	580,855
Age	43.42	9.69	18	90	580,855
Age^2	1978.77	852.04	324	8100	580,855
Tenure	11.18	9.27	0	64	580,855
$Tenure^2$	210.90	304.76	0	4096	580,855
logCompensation	10.89	0.59	9.21	13.82	580,855
logWealth	9.75	1.58	0	14.59	580,855
Equity%	0.78	0.34	0	1	580,855
NoEquity	0.11	0.31	0	1	580,855

Female is an indicator variable denoting whether the employee is female. Age is the employee's age in years. Tenure is the number of years the employee has been employed by the company. LogCompensation is the logarithm of the employee's annual salary. LogWealth is the logarithm of the employee's wealth rating as measured by the IXI value associated with the subject's nine-digit ZIP code. Equity% is the percent of total 2001 contribution that the employee allocated to equity funds. NoEquity is an indicator variable that denotes whether the subject contributed only to money market and bond funds. For all variables, the level of observation is the employee.

Table IV: Descriptive Statistics of the Plans

	Mean	St. Dev.	Min	Max	Obs.
Number of Funds	12.49	6.86	4	59	638
Match	49.44	34.13	0	250	638
Company Stock Offered	0.16	0.37	0	1	638
Restricted Match	3.87	15.97	0	100	638
DB Plan Offered	0.34	0.47	0	1	638
Percent Web Use	0.26	0.13	0	0.91	638
log Number Employees	5.77	1.64	1.10	11.15	638
log Plan Average Compensation	10.98	0.47	9.27	13.39	638
log Plan Average Wealth	10.79	0.79	8.04	14.22	638
Percent Female	0.38	0.19	0	1	638
Plan Average Age	42.82	3.87	30.47	59.67	638
Plan Average Tenure	9.35	4.51	0.27	26.76	638

Number of Funds is the number of funds offered by the plan. Match is the percentage rate at which the employer matches contributions to the plan. Company Stock Offered is an indicator variable denoting whether the plan offered company stock. Restricted Match is an interaction variable equal to zero if employee contributions are only matched with company stock, and equal to Match otherwise. DB Plan offered is an indicator variable denoting whether a defined benefits plan was available to employees. Percent Web Use is the percent of plan participants who registered for online access to their 401(k) accounts. Log Number Employees is the logarithm of the number of people employed at the company. Log Plan Average Compensation is the logarithm of the mean of the employees' yearly salaries. Log Plan Average Wealth is the logarithm of the mean of the employees' wealth ratings, as measured by IXI values for each participant's nine-digit ZIP code. Percent Female is the percentage of employees who are female. Plan Average Age is the mean of the employees' age. Plan Average Tenure is the mean of the number of years the employees have been employed by the company. For all variables, the level of observation is the plan.

Table V: Effect of Number of Funds on Allocation

Dependent Variable	Equity% (1)		NoEquity (2)	
Number of Funds / 100	**-0.328**	**(0.117)****	**0.287**	**(0.104)****
Female	-0.004	(0.004)	-0.006	(0.004)
Age	0.007	(0.001)**	-0.006	(0.001)**
Age^2 / 100	-0.013	(0.001)**	0.010	(0.001)**
Tenure	-0.004	(0.001)**	0.002	(0.001)**
$Tenure^2$ / 100	0.003	(0.003)	-0.003	(0.003)
log Compensation	0.072	(0.007)**	-0.052	(0.007)**
log Wealth	0.013	(0.001)**	-0.008	(0.001)**
Match / 100	0.026	(0.024)	-0.031	(0.027)
Company Stock Offered	0.031	(0.019)	-0.041	(0.021)*
Restricted Match / 100	0.109	(0.034)**	-0.157	(0.038)**
DB Plan Offered	-0.031	(0.018)	0.017	(0.019)
Percent Web Use	-0.057	(0.079)	0.001	(0.076)
log Number Employees	-0.004	(0.005)	0.005	(0.006)
log Plan Average Compensation	-0.005	(0.043)	0.033	(0.043)
log Plan Average Wealth	0.014	(0.021)	-0.024	(0.020)
Percent Female	-0.001	(0.064)	0.005	(0.071)
Plan Average Age	0.004	(0.004)	0.001	(0.004)
Plan Average Tenure	-0.001	(0.004)	0.000	(0.003)
Observations	580,855		580,855	
R^2	0.06		0.05	

(1) Ordinary Least Squares; (2) Linear Probability Model. Robust standard errors in parentheses, clustered by plan. The dependent variable in the first regression is Equity%, which is the percent of total 2001 contribution that a subject allocated to equity funds. The dependent variable in the second regression is NoEquity, an indicator variable that denotes whether the subject contributed only to money market and bond funds. Number of Funds is the number of funds offered by the plan. Female is an indicator variable denoting whether a subject is female. Age is the subject's age in years. Tenure is the number of years the subject has been employed by the company. LogCompensation is the logarithm of the subject's annual salary. LogWealth is the logarithm of the subject's wealth rating as measured by the IXI value associated with the subject's nine-digit ZIP code. Match is the percentage rate at which employers match employee contributions to the plan. Company Stock Offered is an indicator variable denoting whether the plan offered company stock. Restricted Match is an interaction variable equal to zero if employee contributions are only matched with company stock, and equal to Match otherwise. DB Plan offered is an indicator variable denoting whether a defined benefits plan was available to employees. Percent Web Use is the percent of plan participants who registered for online access to their 401(k) accounts. Log Number Employees is the logarithm of the number of people employed at the company. Log Plan Average Compensation is the logarithm of the mean of the employees' yearly salaries. Log Plan Average Wealth is the logarithm of the mean of the employees' wealth ratings, as measured by IXI values for each participant's nine-digit ZIP code. Percent Female is the percentage of employees who are female. Plan Average Age is the mean of the employees' age. Plan Average Tenure is the mean of the number of years the employees have been employed by the company.
* significant at 5%; ** significant at 1%

Table VI: Endogeneity and Selection

Dependent Variable	Equity%		NoEquity		Number of Funds	
	(3)		(4)		(5)	
Female	-0.004	(0.004)	-0.006	(0.004)	0.054	(0.062)
Age	0.007	(0.001)**	-0.006	(0.001)**	0.068	(0.031)*
Age^2 / 100	-0.012	(0.001)**	0.009	(0.001)**	-0.081	(0.036)*
Tenure	-0.004	(0.001)**	0.002	(0.001)**	0.048	(0.028)
$Tenure^2$ / 100	0.004	(0.003)	-0.003	(0.003)	-0.171	(0.092)
logCompensation	0.072	(0.006)**	-0.052	(0.007)**	-0.051	(0.119)
logWealth	0.013	(0.001)**	-0.008	(0.001)**	0.023	(0.034)
Match / 100	0.030	(0.026)	-0.035	(0.028)	-1.339	(1.285)
Company Stock Offered	0.037	(0.020)	-0.046	(0.022)*	-1.737	(1.046)
Restricted Match / 100	0.104	(0.035)**	-0.152	(0.037)**	1.699	(2.259)
DB Plan Offered	-0.033	(0.018)	0.019	(0.020)	0.707	(0.788)
Percent Web Use	-0.108	(0.085)	0.046	(0.080)	15.597	(4.552)**
log Number Employees	-0.003	(0.006)	0.004	(0.006)	-0.275	(0.337)
log Plan Average Compensation	-0.018	(0.043)	0.044	(0.044)	3.848	(1.602)*
log Plan Average Wealth	0.017	(0.021)	-0.027	(0.020)	-1.118	(1.125)
Percent Female	-0.027	(0.073)	0.028	(0.079)	8.068	(3.729)*
Plan Average Age	0.003	(0.004)	0.002	(0.004)	0.315	(0.195)
Plan Average Tenure	-0.001	(0.004)	0.000	(0.003)	-0.025	(0.177)
Observations	580,855		580,855		580,855	
R^2	0.06		0.05		0.17	

(1) Ordinary Least Squares; (2) Llinear Probability Model; (3): Ordinary Least Squares. Robust standard errors in parentheses, clustered by plan. The dependent variable in the first regression is Equity%, which is the percent of total 2001 contribution that a subject allocated to equity funds. The dependent variable in the second regression is NoEquity, an indicator variable that denotes whether the subject contributed only to money market and bond funds. The dependent variable in the third regression is Number of Funds, the number of funds offered by the plan. Female is an indicator variable denoting whether a subject is female. Age is the subject's age in years. Tenure is the number of years the subject has been employed by the company. LogCompensation is the logarithm of the subject's annual salary. LogWealth is the logarithm of the subject's wealth rating as measured by the IXI value associated with the subject's nine-digit ZIP code. Match is the percentage rate at which employers match employee contributions to the plan. Company Stock Offered is an indicator variable denoting whether the plan offered company stock. Restricted Match is an interaction variable equal to zero if employee contributions are only matched with company stock, and equal to Match otherwise. DB Plan offered is an indicator variable denoting whether a defined benefits plan was available to employees. Percent Web Use is the percent of plan participants who registered for online access to their 401(k) accounts. Log Number Employees is the logarithm of the number of people employed at the company. Log Plan Average Compensation is the logarithm of the mean of the employees' yearly salaries. Log Plan Average Wealth is the logarithm of the mean of the employees' wealth ratings, as measured by IXI values for each participant's nine-digit ZIP code. Percent Female is the percentage of employees who are female. Plan Average Age is the mean of the employees' age. Plan Average Tenure is the mean of the number of years the employees have been employed by the company.
* significant at 5%; ** significant at 1%

Filgure 1: Fraction of subjects selecting a gamble as a function of the choice set in Experiment 1

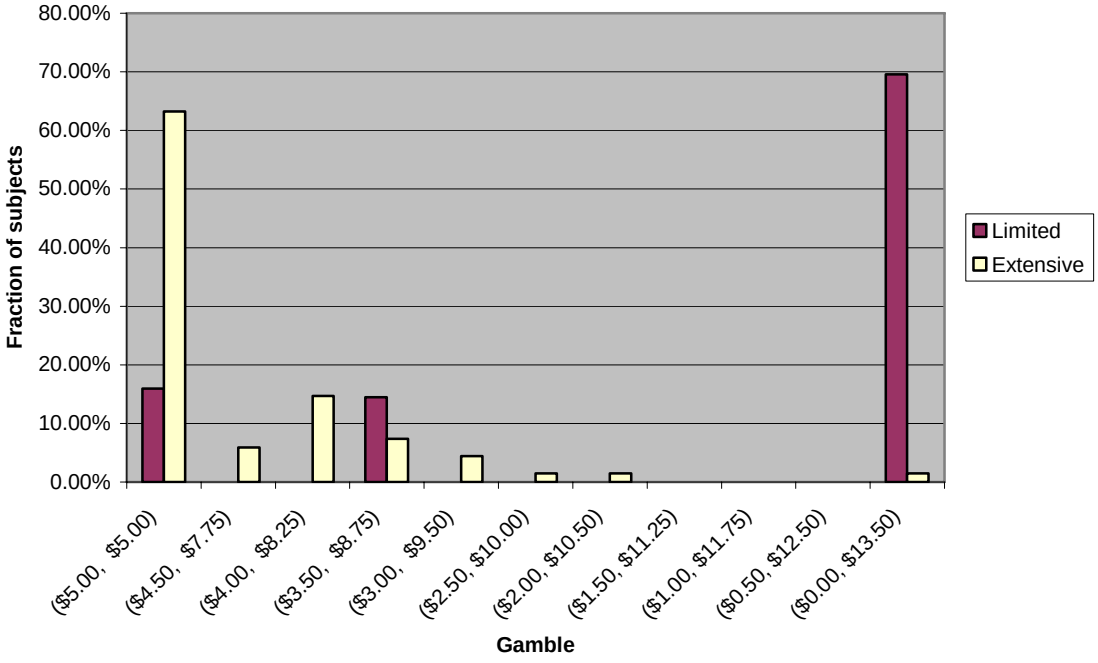
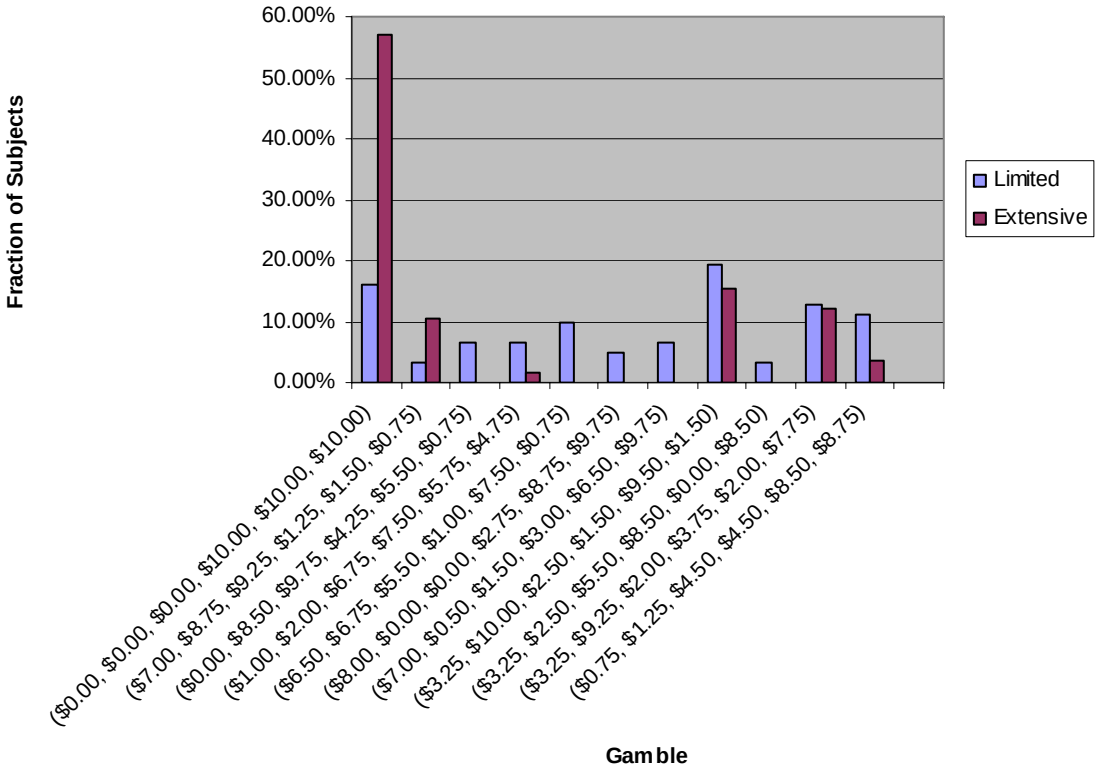


Figure 2: Fraction of subjects selecting a gamble as function of the choice set in Experiment 2



Appendix

Instructions for Experiment 1

We are interested in gathering Columbia students' opinions about other renowned universities. Each of the following 5 pages lists a college or university at the top and a brief list of questions probing your opinion about that college or university. The questions are relatively straightforward. Answer to the best of your ability.

Thank you for participating!

☺

Thank you for participating in the experiment. For compensation, please select one of the gambles below. The experimenter will then flip a coin. Should the coin land on "heads" you will receive the amount specified in the left column. Should the coin land on "tails" you will receive the amount specified in the right column. Please check off the desired gamble and the experimenter will proceed to flip the coin.

Please place a check next to the desired option	If the coin indicates "heads"	If the coin indicates "tails"
	$4.50	$7.75
	$10.50	$2.00
	$2.50	$10.00
	$12.50	$0.50
	$8.75	$3.50
	$1.50	$11.25
	$9.50	$3.00
	$5.00	$5.00
	$11.75	$1.00
	$4.00	$8.25
	$13.50	$0.00

Instructions for Experiment 2

We are interested in gathering Columbia students' opinions about other renowned universities. Each of the following 5 pages lists a college or university at the top and a brief list of questions probing your opinion about that college or university. The questions are relatively straightforward. Answer to the best of your ability.

Thank you for participating!

☺

Thank you for participating in the experiment. For compensation, please select one of the gambles below. The experimenter will provide you with a die. You will cast the die and, depending on how the die falls, receive the amount of money indicated in the table below. Please check off the desired gamble.

Please place a check next to the desired option	If the die falls on 1, you receive	If the die falls on 2, you receive	If the die falls on 3, you receive	If the die falls on 4, you receive	If the die falls on 5, you receive	If the die falls on 6, you receive
	$0.75	$9.25	$8.75	$7.00	$1.25	$1.50
	$0.00	$0.75	$4.25	$5.50	$8.50	$9.75
	$0.00	$0.00	$0.00	$10.00	$10.00	$10.00
	$1.00	$2.00	$6.75	$7.50	$5.75	$4.75
	$1.00	$7.50	$0.75	$6.50	$5.50	$6.75
	$8.00	$0.00	$2.75	$9.75	$0.00	$8.75
	$0.50	$3.00	$1.50	$9.75	$7.00	$6.50
	$2.50	$3.25	$9.50	$1.50	$10.00	$1.50
	$8.50	$3.25	$2.50	$8.50	$0.00	$5.50
	$2.00	$3.25	$3.75	$9.25	$7.75	$2.00
	$4.50	$4.50	$8.75	$8.50	$0.75	$1.25

References

Benartzi, S. and Thaler, R. (2001), "Naïve Diversification Strategies in Defined Contribution Saving Plans," *American Economic Review*, **91**, 79-98.

Bertrand, M., Karlin, D., Mullainathan, S., Shafir, E. and Zinman, J. (2005), "What's Psychology Worth? A Field Experiment in the Consumer Credit Market," National Bureau of Economic Research, Cambridge.

Boatwright, P., and Nunes, J.C. (2001), "Reducing Assortment: An Attribute-Based Approach," *Journal of Marketing*, **65** (3), 50-63.

Choi, J. J., Laibson, D. and Metrick, A. (2002), "How does the Internet Affect Trading? Evidence from Investor Behavior in 401(k) Plans," *Journal of Financial Economics*, **64**, 67-114.

Duflo, E. and Saez, E. (2003), "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics*, **118** (3), 815-42.

Holden, S., and VanDerhei, J. (2003), "401(k) Plan Asset Allocation, Account Balances, and Loan Activity in 2001," Washington, DC: Employee Benefit Research Institute.

Huberman, G., Iyengar, S.S., and Jiang, W. (2004), "Defined Contribution Pension Plans," *mimeo*, Columbia Business School.

Huberman, G. and Jiang, W. (2006), "Offering versus Choice in 401(k) Plans: Equity Exposure and the Number of Funds," *Journal of Finance*, **61** (2), 763-801.

Irons, B., and Hepburn, C. (2006), "Regret Theory and the tyranny of choice," *Economic Record*, forthcoming.

Iyengar, S.S., and Lepper, M. (2000), "When Choice is Demotivating: Can One Desire Too Much of a Good Thing?," *Journal of Personality and Social Psychology*, **79** (6), 995-1006.

Iyengar, S.S., Huberman, G., and Jiang, W. (2004), "How Much Choice is Too Much: Determinants of Individual Contributions in 401K Retirement Plans," in O.S. Mitchell and S. Utkus (Eds.), *Pension Design and Structure: New Lessons from Behavioral Finance* (pp. 83-95). Oxford: Oxford University Press.

Levitt, S.D. and List, J.A. (2006), "What do Laboratory Experiments Tell Us about the Real World?," *mimeo*, University of Chicago.

Kamenica, E. (2006), "Contextual Inference in Markets: On the Informational Content of Product Lines," *mimeo*, University of Chicago.

Kuksov, D., and Villas-Boas, J.M. (2005), "Endogeneity and Individual Consumer Choice," *mimeo*, University of California, Berkeley.

Mottola, G.R., and Utkus, S.P. (2003), "Can There Be Too Much Choice In a Retirement Savings Plan?" The Vanguard Center for Retirement Research, Valley Forge.

Salgado, M. (2005), "Choosing to Have Less Choice," *mimeo*, Northwestern University, Evanston.

Sarver, T. (2005), "Anticipating Regret: Why Fewer Options May Be Better," *mimeo*, Boston University, Boston.

Thaler, R. and Benartzi, S. (2004). "Save More Tomorrow: Using Behavioral Economics to Increase Employee Saving," *Journal of Political Economy*, **112** (1), part 2, S164-S187.

Tversky, A. and Kahneman, D. (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, **5** (4), 297-323.