Journal of Public Economics 94 (2010) 530–539

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Choice proliferation, simplicity seeking, and asset allocation ☆

Sheena S. Iyengar [a], Emir Kamenica [b,*]

[a] Columbia University Graduate School of Business
[b] University of Chicago Booth School of Business

ARTICLE INFO

Article history:
Received 11 November 2008
Received in revised form 22 March 2010
Accepted 30 March 2010
Available online 11 April 2010

Keywords:
Choice overload
Contextual inference
Retirement savings
401(k) plans

ABSTRACT

In settings such as investing for retirement or choosing a drug plan, individuals typically face a large number of options. In this paper, we analyze how the size of the choice set influences which alternative is selected. We present both laboratory experiments and field data that suggest 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 large 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.

© 2010 Elsevier B.V. All rights reserved.

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 purchase a good (Tversky and Shafir, 1992; Iyengar and Lepper, 2000; Boatwright and Nunes, 2001), take up a loan (Bertrand et al., 2010), or enroll in a 401(k) plan (Iyengar et al., 2004), can diminish when the size of the choice set increases. Theoretical investigations show that these instances of *choice overload* can arise when the available choice set conveys useful information (Kamenica, 2008) or when agents exhibit preferences with regret (Irons and Hepburn, 2007; Sarver, 2008). Previous research on choice overload has focused exclusively on how increasing assortment size affects *whether* an agent will participate in a market. In this paper, we investigate how increasing assortment size affects *which* alternative an agent will choose.

We examine this question both in the laboratory and in the field. In the first laboratory experiment, we offer subjects a choice of either 11 or 3 gambles. Specifically, one group of subjects selects a gamble from the menu of 11 gambles which includes 10 risky options and one degenerate gamble ($5 for sure). Another group of subjects is offered a 3-gamble subset of the 11 gambles which includes the degenerate one. We find that many more subjects choose the simple option ($5 for sure) from the set with 11 options than from the set with 3.

In the first experiment, the simplest option is also the least risky. The second experiment removes this confound 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 gambles. We find that, as in the first experiment, the simpler gamble is selected much more often from the larger choice set than from the smaller choice sets.[1]

Then, using data from the Vanguard Center for Retirement Research, we analyze the investment decisions of over 500,000 employees in 638 firms. A context such as investing for retirement is of particular interest for our question since those retirement decisions are of great consequence and the choice sets that employees face are typically quite large. Moreover, the large majority of individuals do save for retirement, which means that the impact of the choice set on which types of options are selected is as important as its impact on participation. Finally, government policy can have a big impact on the number of options that individuals face in this particular context.

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

☆ Early versions of this paper were circulated under the title "Choice overload and simplicity seeking".

* Corresponding author.
E-mail addresses: ss957@columbia.edu (S.S. Iyengar), emir@uchicago.edu (E. Kamenica).

[1] These experimental results are most closely related to Dean's (2008) concurrent finding that larger choices lead to a stronger preference for the default gamble.

0047-2727/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpubeco.2010.03.006

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.

We offer one potential rationalization of our results by noting that, in market settings, small choice sets tend to be better *on average* than large choice sets since small choice sets include only the more select options (Kamenica, 2008). Therefore, a rational uninformed individual might be less likely to select a complex option she does not understand well when many options are available.

The next section reports the results of our randomized experiments. Section 3 describes the 401(k) data and analyzes the relationship between the number of funds in a 401(k) plan and employees' asset allocations. Section 4 discusses the potential rationalization of the results. Section 5 concludes.

2. The experimental tests

Our experiments were conducted at Columbia University. Research assistants 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.

In each of the two main experiments, after completing the survey the subjects were shown a set of gambles from which they chose one as compensation. The subjects assigned to the Extensive conditions were offered a set of 11 gambles, while those in the Limited conditions were offered a subset of 3 gambles out of the original 11. In each of the experiments, one gamble was designed to be the "simple" gamble (e.g., in the first experiment it was $5 for certain instead of a 50–50 chance of two distinct amounts). After the completion of the two main experiments, we conducted an auxiliary study that verified the "simple" gambles were indeed simpler than the others in the sense that it takes less time to compute their expected values.

The order in which the gambles were presented was randomized across subjects. The gambles were constructed so that those with higher expected values have a higher variance.[2] All gambles within an experiment have similar, though not identical, prospect theoretic values as calibrated using estimates from Tversky and Kahneman (1992). The two main experiments differed only in the structure of the gambles. Sample instructions (excluding parts of the unrelated questionnaire) are provided in Appendix A.

2.1. Experiment 1

In the first experiment, subjects chose from 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,' she gets $7.75), and one was a sure bet ($5.00 whether the coin falls on 'heads' or 'tails'). We will refer to this degenerate gamble as the simple gamble.[3] The 3 gambles presented to the subjects in the Limited condition included the simple one.[4] Table 1 provides the lists of the gambles in the two conditions. Note that, as the instructions in Appendix A reveal, the simple gamble was presented in exactly the same format as the other gambles and was embedded in the list of the other gambles.

Table 1
Set of gambles for Experiment 1.

Gamble #	If heads	If tails
Extensive condition		
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		
1	$5.00	$5.00
2	$3.50	$8.75
3	$0.00	$13.50

2.1.1. Results

We observe a dramatic 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). In other words, subjects were roughly four times more likely to select the $5 for sure when facing 10 other options than when facing only 2 other options. Fig. 1 depicts the histogram of the distribution of choices in the two conditions.

2.2. Experiment 2

In the first experiment, the simplest option was also the least risky. Experiment 2 is designed to test the hypothesis that simplicity, rather than lower risk, becomes more attractive as the size of the choice set grows. In the second experiment, there were six possible outcomes associated with each gamble, and compensation was determined by a die toss. Of the 11 gambles, 10 gambles yielded a distinct amount[5] (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 gamble was riskier, paying out either $0 (on a ⚀, ⚁, or ⚂) or $10 (on a ⚃, ⚄, or ⚅). We will refer to this all-or-nothing gamble as the simple gamble.[6] The subset of 3 gambles presented to the subjects in the Limited condition always included the simple gamble and two others which were randomly selected, independently for each subject, from the other 10 gambles in the Extensive condition. As in the previous experiment, the simple gamble was presented in the same manner as the others. Table 2 provides the list of gambles in the Extensive condition.

[2] This relationship is not strict. 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.

[3] In the auxiliary experiment we incentivized 54 Columbia undergraduates to mentally compute the expected values of the gambles. Conditional on giving a correct answer, the median time it took to calculate the expected value of the simple gamble was 4.87 s, while the median times it took to calculate the expected value of the other 10 gambles ranged from 7.59 to 16.12 s. We did not find that subjects were more likely to give the correct answer when computing the expected value of the simple gamble.

[4] The other two gambles were selected at random 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 select the two additional gambles independently for each subject.

[5] 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.

[6] In the auxiliary experiment we incentivized 54 Columbia undergraduates to mentally compute the expected values of the gambles. Conditional on giving a correct answer, the median time it took to calculate the expected value of the simple gamble was 6.34 s, while the median times it took to calculate the expected value of the other 10 gambles ranged from 18.70 to 60.57 s. Moreover, subjects were more likely to give the correct answer when computing the expected value of the simple gamble (94% versus 68–80% for the other gambles).

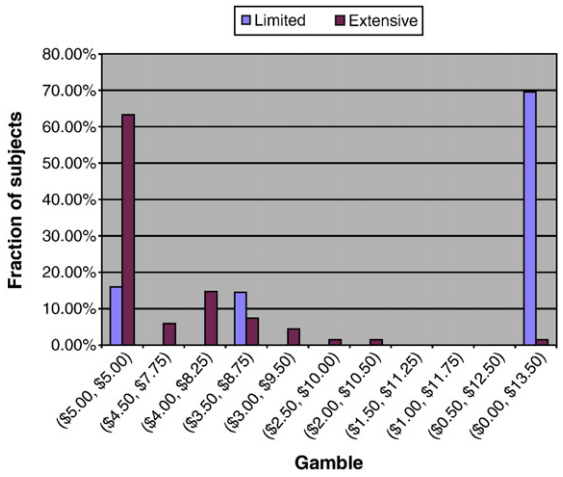


Fig. 1. Fraction of subjects selecting a gamble as a function of the choice set in Experiment 1.

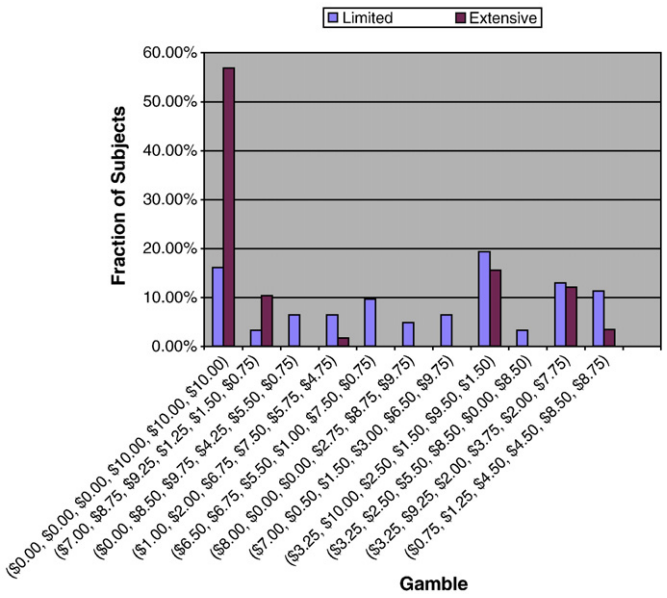


Fig. 2. Fraction of subjects selecting a gamble as function of the choice set in Experiment 2.

2.2.1. Results

Once again, the simple 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). Fig. 2 illustrates the histogram of the distribution of choices in the two conditions.

These experimental findings establish that subjects have a stronger preference for the simplest option when choosing from a larger choice set. The next section examines how the size of the choice sets affects asset allocation decisions in 401(k) plans. This setting is clearly of much greater policy relevance, but has the limitation that we have less evidence on what constitutes a simple option when it comes to 401(k) funds.

3. Fund availability and asset allocation

401(k) plans, employer-sponsored plans in which employees are given monetary incentives to transfer some of their salary into investment funds provided by the plan, are an increasingly common vehicle for retirement savings. As of year-end 2008, 50 million American workers held $2.3 trillion in assets in their 401(k) plans (Holden et al., 2009).

The features of a 401(k) plan, such as the choice of the menu of funds to be offered, are determined jointly by the employer and the company administering the plan, such as Vanguard. Given that a substantial share of all assets for retirement are held in 401(k) plans, the question of whether the resulting design of an individual's plan affects his or her investment behavior attains great importance. The existing literature on this question is divided.

On one hand, a number of papers, recently summarized by Benartzi and Thaler (2007), suggest that the design of the 401(k) plan influences savings behavior in various ways. For example, Madrian and Shea (2001) report that savings rates and allocations are sensitive to default options, Benartzi and Thaler (2001) and Brown et al. (2007) argue that the menu of funds offered has a strong effect on portfolio choices, and Benartzi and Thaler (2002) demonstrate that a particular presentation of investment portfolios leads employees to prefer the median portfolio in their institution to the one they had chosen for themselves. Iyengar et al. (2004) report that a greater number of funds in a 401(k) plan are associated with lower participation rates. Beshears et al. (2006) show that simplifying the participation decision by providing an easy way to select a particular contribution rate and allocation substantially increases participation rates. More broadly, Benartzi et al. (forthcoming) argue that even minor details of the retirement plan design can have dramatic effects on savings behavior.

On the other hand, Huberman and Jiang (2006) contest the sensitivity of asset allocation to the features of the 401(k) plan and report that: (i) the composition of funds offered (e.g., fraction of funds that are equity funds) does not sizeably affect asset allocation, and (ii) the number of funds offered does not affect the number of funds used by a plan participant.[7]

Our paper contributes to this debate by examining how the size of the choice set affects people's behavior in context of 401(k) plans: we ask whether, conditional on participation, the number of funds offered influences asset allocation.

Table 2
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

3.1. The data

The data was provided by the Vanguard Center for Retirement Research, whose records from 2001 include 639 defined contribution (DC) pension plans[8] with 588,926 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 is 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

[7] Benartzi and Thaler (2007) argue that this seeming lack of naïve diversification stems only from the design of the sign-up form.

[8] For the purpose of this analysis, a plan refers to an institution which provides 401(k) funds to eligible employees.

investment advice (Mottola and Utkus, 2003).[9] Moreover, none of the plans in our data provided a default option for asset allocation: each individual had to make an active decision (Carroll et al., 2009) about their allocation.

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*.[10] Employees' wealth is measured through their IXI index: a company called IXI collects retail and IRA asset data from large financial services companies at the 9-digit Zip Code level. IXI then assigns a wealth rank (from 1 to 24) to the Zip Codes, based on the imputed average household assets.

We exclude individuals whose income is below $10,000 or above $1,000,000, and 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 a fund category. These criteria leave us with 580,855 employees in 638 distinct plans.

Our dependent variables are measures of the employees' contributions across different types of funds. The data provides information on how individuals allocated their total annual 401(k) contribution in 2001 (including the employer match) across 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 categories of funds that contain equities (active funds, balanced funds, company stock, and index funds). Similarly, *Bond%* and *Cash%* indicate what fraction of the total 2001 contribution an employee placed in pure bond funds and money market funds, respectively. While we arbitrarily categorize balanced funds as equity, any other categorization (e.g., counting them as 1/2 bond and 1/2 equity, or counting them fully as bond funds) yields similar (and slightly stronger) results. We exclude "other" funds from the analysis, but since only 96 employees allocated a positive amount to these funds, any categorization of "other" leaves the results unchanged. We also define an indicator variable *NoEquity*, which takes the value of 1 if an employee allocated her entire contribution only to money market and bond funds, and value 0 otherwise. These four variables: *Equity%*, *Bond%*, *Cash%* and *NoEquity* will be our measures of employees' asset allocation.[11] Table 3 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 offer some match to their employees' contributions. Variable *Match* indicates the employer's match rate. In our sample, 102 plans have "own-company stock" as an investment option. Variable *CompanyStockOffered* is an indicator variable for whether the plan offers company stock. *RestrictedMatch* is an interaction variable equal to zero if the match is with company stock only and equal to *Match* otherwise. A defined benefits (DB) 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. Indicator variable *DBPlanOffered* indicates whether a defined benefits plan is available. Almost all of the plans offer internet accessibility. Variable *PercentWebUse* represents the percent of plan participants who registered for web access to their 401(k) accounts. Variable *LogNumberEmployees* captures the size of the firm.

Table 3
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
Bond%	0.06	0.17	0	1	580,855
Cash%	0.16	0.31	0	1	580,855
NoEquity	0.11	0.31	0	1	580,855

Female is an indicator variable denoting whether the employee is female, or if that individual's information is missing, the percentage of female employees in the plan. 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 containing funds. Bond% is the percent of total 2001 contribution that the employee allocated to bond funds. Cash% is the percent of total 2001 contribution that the employee allocated to money market 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.

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 might 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 number of funds available in a plan varies from 4 to 59.[12] The mean and median number of funds in a plan are 12.49 and 11, respectively. The standard deviation around the mean is 6.86. Ninety percent of plans offer between 6 and 20 fund choices, and 18 plans offer 30 options or more. Table 4 provides summary statistics at the plan level and Fig. 3 depicts the distribution of *NumberOfFunds*.

3.2. The impact of number of funds on allocation

We first consider OLS regressions of the form:

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

where $Category\%_{ij}$ is the percentage of contribution individual i in plan j allocated to a particular category (*Equity*, *Bond*, or *Cash*), X_{ij} is the vector of individual-specific attributes: {*Female*, *Age*, *Age*^{2}, *Tenure*, *Tenure*^{2}, *LogCompensation*, *LogWealth*}, and Z_j denotes plan-level characteristics: {*Match*, *CompanyStockOffered*, *RestrictedMatch*, *DBPlanOffered*, *PercentWebUse*, *LogNumberEmployees*, *LogPlanAverageCompensation*, *LogPlanAverageWealth*, *PercentFemale*, *PlanAverageAge*, *PlanAverageTenure*}. We cluster the standard errors by plan.

Column (1) of Table 5 reports the impact of the number of funds on the contribution to equities. On average, for every 10 funds added to a plan, the allocation to equity funds decreases by 3.28 percentage points ($p = 0.005$).[13] Fig. 4 plots the residuals of the regression of *Equity%* on the individual and plan-level controls against the percentile of the residuals of the regression of *NumberOfFunds* on those controls. The scatter plot suggests that the negative relationship

[9] 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.

[10] 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.

[11] Unfortunately, we do not observe any asset holdings that the employees might have outside their 401(k) accounts. As with most other studies of retirement savings behavior, there is little we can do to overcome this limitation of the data.

[12] In the raw data, for one single observation, the number of funds was apparently miscoded as 2. Under the employee inclusion criteria specified earlier we exclude that observation from our analysis.

[13] Recall that the standard deviation of the number of funds is around 7.

Table 4
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.

between the number of funds and allocation to equity holds throughout the range of the data. Given the mean allocation to equity funds of 78%, however, the decrease of 3.28 percentage points for every 10 funds is probably of limited economic significance.

Column (2) reveals that some of the decrease in exposure to equities stems from an increased allocation to bond funds. On average, for every 10 funds added to a plan, the allocation to bond funds increases by 1.98 percentage points ($p=0.006$), relative to the mean of 6%. Since *Equity%*, *Bond%*, and *Cash%* must add up to one, these coefficients in columns (1) and (2) imply that the exposure to money market funds increases by 1.30 percentage points for every 10 additional funds in the plan, as indicated in column (3), though this effect is not statistically significant ($p=0.244$).

Finally, column (4) examines the impact of the number of funds on the probability that an employee invests no money whatsoever in equity funds, using a linear probability model:

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

While on average only 10.53% of employees do not invest any money in equities, this probability increases by 2.87 percentage points, or around 27%, for every 10 additional funds ($p=0.006$).[14] Fig. 5 plots the residuals of the regression of *NoEquity* on the individual and plan-level controls against the percentile of the residuals of the regression of *NumberOfFunds* on those controls. The scatter plot again suggests that the negative relationship between participation in equity and the number of funds holds throughout the range of the data, though may be somewhat less strong in the upper half of the distribution of the residuals of *NumberOfFunds*.

[14] One might suspect that the design of the 401(k) plan might have a greater effect on asset allocation for those employees who are more likely to be unfamiliar with asset allocation decisions. Accordingly, we considered specifications that include an interaction term between the number of funds and proxies for employee sophistication, namely income and wealth, and specifications that estimate the models (1) and (2) separately for employees below and above the median on these measures. We were unable to detect any statistically significant heterogeneity in the effect of the number of funds on equity exposure.

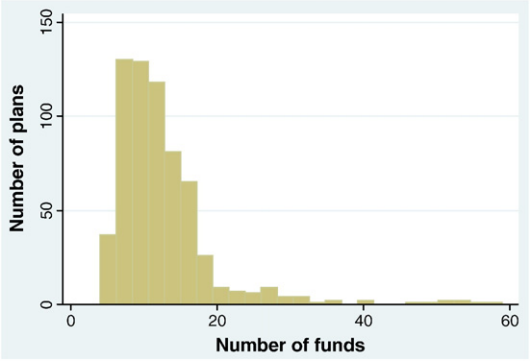


Fig. 3. Distribution of number of funds across plans.

Given that non-participation in the stock market, especially for younger employees, is likely to be detrimental to one's retirement income, this effect is potentially of substantial economic significance. Non-participation is particularly a concern because, in our data, employees under 30 years of age are as likely as others to allocate no money at all to equity funds and their participation in equities is just as sensitive to the number of funds as that of older employees.[15]

Three factors, however, could bring into question the magnitude of the welfare consequences from non-participation. One is that the increase in the number of funds may increase non-participation only by shifting to zero those who would have otherwise invested very little in equities. For example, it might be that the increase in non-participants comes only from those who were investing less than 3% of their assets in equities. A closer look at the data, however, reveals that this is not the case. Even though an employee at the 10.5th percentile of the distribution of equity exposure holds no equities, an employee at the 11th percentile already allocates over 10% of her contribution to equity funds, while an employee at the 13th percentile allocates 25% of her contribution to equity funds. By the 18th percentile, employees invest over half of their contributions to equities. Therefore, even in the most extreme case where marginal non-participants are drawn exclusively from the bottom of the distribution of equity exposure, the drop to non-participation induced by a plan design with more numerous fund options involves a substantial decrease in equity exposure.

The second concern which might diminish the economic importance of our estimates is that we only observe individuals' allocation to their 401(k) plans; if people in our sample hold substantial assets elsewhere, then their lack of equity exposure in the 401(k) plan might imply no loss of welfare. This may be a particularly important concern because asset location considerations (e.g., Shoven and Sialm, 2003) suggest that households should hold their highly taxed fixed-income securities in their retirement accounts (such as a 401(k) plan) and place their less-highly taxed equity securities in their taxable accounts. An examination of the correlates of equity exposure in our sample, however, strongly mitigates the possibility that asset location plays an important role in our findings. As Table 7 reveals, individuals with higher income and higher wealth are far more likely to invest in equities, which is contrary to what we would expect if considerations of asset location were an important factor.

Finally, we consider the possibility that non-participants would have limited gains from participating since they would be less sophisticated investors and would invest inefficiently (Calvet et al., 2007). While Calvet et al. (2007) calculate that this might reduce the costs of non-participation by as much as one-half, even 50% of the potential loss in utility from non-participation in equities involves a

[15] For employees under 30, the coefficient of *NoEquity* on *NumberOfFunds* is actually greater (0.361) than for older employees (0.282), but this difference is not statistically significant.

Table 5
Effect of the number of funds on allocation.

Dependent variable	Equity%		Bond%		Cash%		NoEquity	
	(1)		(2)		(3)		(4)	
Number of Funds/100	**−0.328**	**(0.117)****	**0.198**	**(0.072)****	**0.130**	**(0.112)**	**0.287**	**(0.104)****
Female	−0.004	(0.004)	0.007	(0.002)**	−0.003	(0.003)	−0.006	(0.004)
Age	0.007	(0.001)**	0.001	(0.000)	−0.008	(0.001)**	−0.006	(0.001)**
Age^2/100	−0.013	(0.001)**	0.000	(0.001)	0.013	(0.001)**	0.010	(0.001)**
Tenure	−0.004	(0.001)**	0.000	(0.001)	0.003	(0.001)**	0.002	(0.001)**
Tenure^2/100	0.003	(0.003)	−0.002	(0.002)	−0.001	(0.003)	−0.003	(0.003)
Log Compensation	0.072	(0.007)**	−0.006	(0.002)**	−0.066	(0.007)**	−0.052	(0.007)**
Log Wealth	0.013	(0.001)**	0.000	(0.001)	−0.013	(0.001)**	−0.008	(0.001)*,
Match/100	0.026	(0.024)	−0.034	(0.023)	0.008	(0.029)	−0.031	(0.027)
Company Stock Offered	0.031	(0.019)	−0.056	(0.016)**	0.026	(0.020)	−0.041	(0.021)*
Restricted Match/100	0.109	(0.034)**	−0.003	(0.015)	−0.106	(0.034)**	−0.157	(0.038)**
DB Plan Offered	−0.031	(0.018)	0.012	(0.009)	0.019	(0.017)	0.017	(0.019)
Percent Web Use	−0.057	(0.079)	−0.020	(0.037)	0.077	(0.077)	0.001	(0.076)
Log Number Employees	−0.004	(0.005)	0.012	(0.006)*	−0.008	(0.004)	0.005	(0.006)
Log Plan Average Compensation	−0.005	(0.043)	−0.017	(0.017)	0.022	(0.041)	0.033	(0.043)
Log Plan Average Wealth	0.014	(0.021)	0.005	(0.008)	−0.018	(0.020)	−0.024	(0.020)
Percent Female	−0.001	(0.064)	0.198	(0.069)**	−0.197	(0.061)**	0.005	(0.071)
Plan Average Age	0.004	(0.004)	0.005	(0.003)	−0.009	(0.004)*	0.001	(0.004)
Plan Average Tenure	−0.001	(0.004)	0.001	(0.001)	0.000	(0.003)	0.000	(0.003)
Observations	580,855		580,855		580,855		580,855	
R^2	0.06		0.12		0.06		0.05	

(1–3) Ordinary Least Squares; (2) Linear Probability Model. Robust standard errors in parentheses, clustered by plan. Equity%, Bond%, and Cash% are the fraction of allocation the employee put in all equity containing funds, bond funds, and money market funds, respectively. NoEquity is 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. Log Compensation is the logarithm of the subject's annual salary. Log Wealth 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%.

substantial decrease in welfare under any reasonable assumptions about risk aversion. For a rough sense of this magnitude consider a 45-year-old[16] individual with CRRA utility over wealth at retirement at age 65, with the coefficient of relative risk aversion equal to 5.[17] If this person annually invests $3000[18] in a riskless asset with a 2% real return instead of investing that money in a risky asset with a lognormal distribution of returns with the mean of 9% and a standard deviation of 16% (which approximates the historic returns on U.S. stocks), the resulting loss is equivalent to a $17,000 decrease in the individual's current wealth.[19]

The fact that equity exposure and participation fall with the number of funds in a plan is all the more striking because the percentage of funds that are equity funds *increases* in the overall number of funds: for the median plan, roughly 3/4 of the funds are equity funds and this percentage increases by 3.94 percentage points for every 10 additional funds. Hence, as Table 6 shows, both the fraction of contributions allocated to equity funds and the probability that at least some money is allocated to equity funds are decreasing in the number of *equity* funds.[20]

3.3. The interpretation of the 401(k) results

In the context of our experimental gambles, it is apparent that the options whose popularity increased in larger choice sets were simple in the sense in that it is easy to compute their expected values.[21] In the context of 401(k) plans, however, it is more difficult to identify why equities in particular become less popular in larger choice sets.

One possibility is that larger choice sets lead employees to invest in fund categories that they are familiar with. The John Hancock Survey of Defined Contribution Plan Participants asked a nationwide random sample of 801 defined plan participants about their familiarity with various categories of funds.[22] Participants were most familiar with company stock funds, and were more familiar with stock funds than with bond funds. Exposure to bond funds increases, however, and exposure to stock funds decreases when the number of funds is greater. Moreover, we find that exposure to company stock also decreases when there is a greater number of funds.[23] Hence, our results are not consistent with the interpretation that a larger number of funds leads to greater investment in familiar fund categories.

An alternative possibility is that funds are presented to employees in order of increasing risk and the presence of a larger number of funds increases the chance that employees select funds in categories listed at the beginning. Unfortunately, we do not have the data on the order in which the categories are listed. Anecdotally, however,

[16] The median age of the employees in our sample is 44.

[17] We consider this to be the upper end of plausible values for the coefficient of relative risk aversion.

[18] Median annual contribution for individuals in our sample who are 45 years and older is $2984.

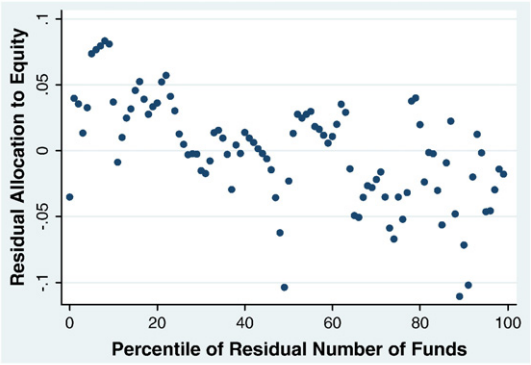
[19] Vissing-Jorgensen (2003) finds much smaller costs of non-participation in equities, with a $55 annual cost being sufficient to explain the non-participation of half the non-participants. The reason why the cost is low in her sample, however, is the low or nonexistent financial wealth of most households. By contrast, we look at employees who *are* contributing to their retirement savings each year and consider the cost of not investing any of this contribution in equities.

[20] Recall that Huberman and Jiang (2006) find that exposure to equities is not sensitive to the *fraction* of equity funds.

[21] As we mentioned in footnotes 3 and 6, we also confirmed this experimentally at the request of a referee.

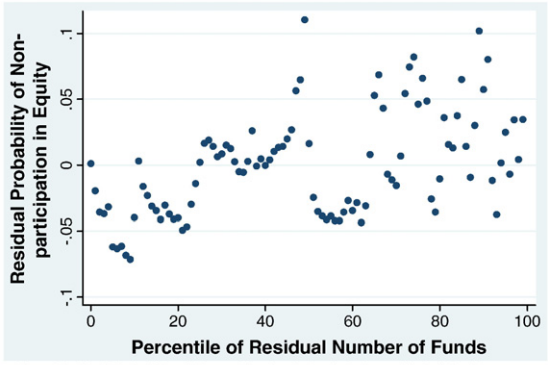
[22] http://www.jhancockstructures.com/gsfp/survey2002.pdf; accessed on January 20, 2010.

[23] Specifically, for every 10 additional funds allocation to company stock decreases by 6 percentage points.



Note: Residual Allocation to Equity is the residual of the OLS regression of *Equity%* on the individual- and plan-level controls. Percentile of Residual Number of Funds is the percentile of the residual of the OLS regression of *NumberOfFunds* on the individual- and plan-level controls.

Fig. 4. Allocation to equity and the number of funds.



Note: Residual Probability of Nonparticipation in Equity is the residual of the Linear Probability Model regression of *NoEquity* on the individual- and plan-level controls. Percentile of Residual Number of Funds is the percentile of the residual of the OLS regression of *NumberOfFunds* on the individual- and plan-level controls.

Fig. 5. Non-participation in equity and the number of funds.

company stock is often listed first, with other categories listed subsequently in order of increasing risk. If this practice is prevalent in the plans in our sample, the finding that allocation to company stock decreases with the number of funds is evidence against order effects playing an important role in our results.[24]

Given that the two aforementioned possibilities are not consistent with the data, the interpretation of the 401(k) results is less straightforward than that of our experimental findings. Nonetheless, the analysis in this section suggests that the size of the choice set can have an impact on important real world decisions, such as allocation of assets in retirement savings.

3.4. Robustness

Our data provides information on the set of funds available in the plans as of 2001 and the employees' contribution to various fund categories in that year. Previous research (Ameriks and Zeldes, 2004) suggests, however, that most people change the allocation of their 401(k) contributions very rarely.[25] Hence, if the set of funds within a plan changed substantially prior to 2001, the variable *NumberOfFunds* is a noisy measure of the number of options that the employees actually faced at the time when they made their decision. To address this concern, we examine whether our results change when we focus only on the employees that recently joined the plan. When we run specifications (1) and (2) on the subsample of employees that have been in their institution for no more than a year and a half,[26] we find that for every 10 additional funds, allocation to equity funds decreases by 2.78 percentage points ($p=0.050$) while the probability that an employee puts no money whatsoever in equity funds increases by 3.15 percentage points ($p=0.029$). The size of these effects is statistically indistinguishable from the impact of *NumberOfFunds* on allocation of the other 90% of employees.

Our results may also be compromised by selection: plans with different number of funds may have different employees, or, within an institution, the type of employee that self-selects into participation may vary with the number of funds in the plan (Iyengar et al., 2004). To address these two concerns, we compare the way in which *Category%* and *NoEquity* vary with the individual- and plan-level attributes with the way in which *NumberOfFunds* does so.[27] Specifically, we compare the coefficients β_1 and β_2 across these OLS regressions:

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

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

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

Table 7 reports the results. As one might expect, 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 of funds affect the two in the same direction.[28] In a similar vein, Table 8 shows that the coefficients on *NumberOfFunds* in regressions (1) and (2) are not very sensitive to the inclusion of controls. Adding individual- and plan-level covariates either has no effect or somewhat strengthens our results. Therefore, to the extent that the observable characteristics in our data are representative of the unobservables, the omitted variable bias is unlikely to drive our results (Altonji et al., 2005).[29]

4. Discussion

Although our data is not well suited for identifying the mechanism through which the size of the choice set affects which option is selected, we offer one way to reconcile the observed patterns with rational decision making. 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 (Kamenica, 2008). This means that, even though the *best* option becomes better as the choice set becomes larger, the *average* option becomes worse. A decision maker, therefore, may be willing to select a complex option she does not fully understand when only a few options are available (since all elements

[24] We conducted an additional experiment asking subjects to respond to a hypothetical question about how they would allocate their 401(k) savings while varying the number and the order of the fund categories. The responses to this hypothetical question were quite noisy. We detected no significant relationship between the allocation to a given category and its order in the list, nor between the allocation to a given category and the overall number of funds in the list.

[25] Outside of 401(k) plans, however, Calvet et al. (2009) find strong evidence for active rebalancing.

[26] These employees constitute 10.10% of the overall sample.

[27] Given that the number of funds impacts participation only by affecting the speed at which employees join a 401(k) plan (Iyengar et al., 2004), the fact that our results hold for all levels of tenure provides additional evidence that within-plan selection is not an important factor.

[28] The coefficients in Table 7 suggest that unobserved employee sophistication in particular is likely to be associated with more funds and with greater equity exposure.

[29] 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 the other employees in the firm, but unfortunately the relationship between aggregate characteristics of the plan and its number of funds is too weak for this approach.

Table 6
Effect of the Number of Equity Funds on Allocation.

Dependent variable	Equity%		Bond%		Cash%		NoEquity	
	(1)		(2)		(3)		(4)	
Number of Equity Funds/100	**−0.346**	**(0.130)****	**0.282**	**(0.102)****	**0.064**	**(0.117)**	**0.346**	**(0.123)****
Female	−0.004	(0.004)	0.007	(0.002)**	−0.003	(0.003)	−0.006	(0.003)
Age	0.007	(0.001)**	0.001	(0.000)	−0.008	(0.001)**	−0.006	(0.001)**
Age^2/100	−0.013	(0.001)**	0.000	(0.001)	0.013	(0.001)**	0.010	(0.001)**
Tenure	−0.004	(0.001)**	0.000	(0.001)	0.003	(0.001)**	0.002	(0.001)**
Tenure^2/100	0.003	(0.003)	−0.002	(0.002)	−0.001	(0.003)	−0.003	(0.003)
Log Compensation	0.072	(0.006)**	−0.006	(0.002)**	−0.066	(0.007)**	−0.052	(0.007)**
Log Wealth	0.013	(0.001)**	0.000	(0.001)	−0.013	(0.001)**	−0.008	(0.001)**
Match/100	0.027	(0.025)	−0.034	(0.022)	0.007	(0.030)	−0.031	(0.027)
Company Stock Offered	0.032	(0.019)	−0.056	(0.015)**	0.024	(0.020)	−0.041	(0.020)*
Restricted Match/100	0.106	(0.034)**	−0.002	(0.015)	−0.105	(0.034)**	−0.155	(0.038)**
DB Plan Offered	−0.031	(0.018)	0.012	(0.009)	0.019	(0.017)	0.017	(0.019)
Percent Web Use	−0.065	(0.081)	−0.024	(0.039)	0.089	(0.080)	0.003	(0.076)
Log Number Employees	−0.004	(0.005)	0.012	(0.006)*	−0.008	(0.004)	0.005	(0.006)
Log Plan Average Compensation	−0.007	(0.043)	−0.018	(0.017)	0.025	(0.042)	0.033	(0.044)
Log Plan Average Wealth	0.014	(0.021)	0.005	(0.008)	−0.019	(0.020)	−0.024	(0.020)
Percent Female	0.000	(0.064)	0.191	(0.065)**	−0.191	(0.060)**	0.001	(0.069)
Plan Average Age	0.004	(0.004)	0.005	(0.003)	−0.009	(0.004)*	0.000	(0.004)
Plan Average Tenure	0.000	(0.004)	0.001	(0.001)	0.000	(0.003)	0.000	(0.003)
Observations	580,855		580,855		580,855		580,855	
R^2	0.06		0.12		0.06		0.05	

(1–3) Ordinary Least Squares; (4) Linear Probability Model. Robust standard errors in parentheses, clustered by plan. Equity%, Bond%, and Cash% are the fraction of allocation the employee put in all equity containing funds, bond funds, and money market funds, respectively. NoEquity is an indicator variable that denotes whether the subject contributed only to money market and bond funds. Number of Equities is the number of equities 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. Log Compensation is the logarithm of the subject's annual salary. Log Wealth 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 7
Endogeneity and selection.

Dependent variable	Equity%		Bond%		Cash%		NoEquity		Number of Funds	
	(1)		(2)		(3)		(4)		(5)	
Female	−0.004	(0.004)	0.007	(0.003)**	−0.003	(0.003)	−0.006	(0.004)	0.054	(0.062)
Age	0.007	(0.001)**	0.001	(0.000)	−0.008	(0.001)**	−0.006	(0.001)**	0.068	(0.031)*
Age^2/100	−0.012	(0.001)**	0.000	(0.001)	0.013	(0.001)**	0.009	(0.001)**	−0.081	(0.036)*
Tenure	−0.004	(0.001)**	0.000	(0.001)	0.003	(0.001)**	0.002	(0.001)**	0.048	(0.028)
Tenure^2/100	0.004	(0.003)	−0.003	(0.002)	−0.001	(0.003)	−0.003	(0.003)	−0.171	(0.092)
Log Compensation	0.072	(0.006)**	−0.006	(0.002)**	−0.066	(0.007)**	−0.052	(0.007)**	−0.051	(0.119)
Log Wealth	0.013	(0.001)**	0.000	(0.001)	−0.013	(0.001)**	−0.008	(0.001)**	0.023	(0.034)
Match/100	0.030	(0.026)	−0.037	(0.024)	0.006	(0.030)	−0.035	(0.028)	−1.339	(1.285)
Company Stock Offered	0.037	(0.020)	−0.060	(0.017)**	0.023	(0.020)	−0.046	(0.022)*	−1.737	(1.046)
Restricted Match/100	0.104	(0.035)**	0.001	(0.016)	−0.104	(0.034)**	−0.152	(0.037)**	1.699	(2.259)
DB Plan Offered	−0.033	(0.018)	0.014	(0.009)	0.020	(0.017)	0.019	(0.020)	0.707	(0.788)
Percent Web Use	−0.108	(0.085)	0.011	(0.039)	0.097	(0.080)	0.046	(0.080)	15.597	(4.552)**
Log Number Employees	−0.003	(0.006)	0.011	(0.006)	−0.008	(0.004)	0.004	(0.006)	−0.275	(0.337)
Log Plan Average Comp.	−0.018	(0.043)	−0.009	(0.018)	0.027	(0.042)	0.044	(0.044)	3.848	(1.602)*
Log Plan Average Wealth	0.017	(0.021)	0.002	(0.009)	−0.020	(0.019)	−0.027	(0.020)	−1.118	(1.125)
Percent Female	−0.027	(0.073)	0.214	(0.077)**	−0.186	(0.061)**	0.028	(0.079)	8.068	(3.729)*
Plan Average Age	0.003	(0.004)	0.006	(0.003)	−0.009	(0.004)*	0.002	(0.004)	0.315	(0.195)
Plan Average Tenure	−0.001	(0.004)	0.001	(0.002)	0.000	(0.004)	0.000	(0.003)	−0.025	(0.177)
Observations	580,855		580,855		580,855		580,855		580,855	
R^2	0.06		0.12		0.06		0.05		0.17	

(1,2,3,5) Ordinary Least Squares; (4) Linear Probability Model. Robust standard errors in parentheses, clustered by plan. Equity%, Bond%, and Cash% are the fraction of allocation the employee put in all equity containing funds, bond funds, and money market funds, respectively. NoEquity is 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. Log Compensation is the logarithm of the subject's annual salary. Log Wealth 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 8
The impact of controls.

Dependent variable	Equity%				NoEquity			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Funds/100	−0.216	−0.343	−0.310	−0.328	0.236	0.328	0.276	0.287
	(0.164)	(0.159)*	(0.113)**	(0.117)**	(0.180)	(0.173)	(0.102)**	(0.104)**
Individual-level Controls	No	Yes	No	Yes	No	Yes	No	Yes
Plan-level Controls	No	No	Yes	Yes	No	No	Yes	Yes
Observations	580,855	580,855	580,855	580,855	580,855	580,855	580,855	580,855
R^2	0.00	0.05	0.02	0.06	0.00	0.03	0.04	0.05

(1–4) Ordinary Least Squares; (5–8) Linear Probability Model. Robust standard errors in parentheses, clustered by plan. Equity% is the fraction of allocation the employee put in all equity containing funds. NoEquity is 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. Individual-level Controls are *Female*, *Age*, *Age*^{2}, *Tenure*, *Tenure*^{2}, *LogCompensation*, and *LogWealth*. Plan-level Controls are *Match*, *CompanyStockOffered*, *RestrictedMatch*, *DBPlanOffered*, *PercentWebUse*, *LogNumberEmployees*, *LogPlanAverageCompensation*, *LogPlanAverageWealth*, *PercentFemale*, *PlanAverageAge*, and *PlanAverageTenure*.
* Significant at 5%.
** Significant at 1%.

of a small choice set yield high expected utility), but would shy away from such an option when there are many alternatives (since the large choice set also includes niche products that yield low average utility). In other words, to avoid the possibility of selecting an inappropriate niche product, an uninformed individual is better off selecting a simple option when choosing from a large choice set.

In order for this argument to provide a rationalization of our results, we need to make two additional assumptions. The first one is that people believe that choice sets they encounter across a variety of decision contexts share the aforementioned feature of equilibrium product lines. In the decision contexts we examine, namely the experimental setting and the 401(k) allocation decision, rationality *per se* does not pin down the subjects' and employees' beliefs about the structure of the choice sets since the incentives of the experimenters and the 401(k) plan designers are not obvious. However, it seems reasonable that in these settings people might act as if the choice set they face is not unlike the everyday choice sets they encounter in the market.

The second assumption is that uninformed 401(k) participants interpret the category of equity funds as potentially containing more niche options. We do not have information on the particular funds offered in the plans in our sample, but the 401(k) plan offered by the home institutions of the authors (also managed by Vanguard), includes numerous specialty equity funds (e.g., energy, health care, precious metals and mining, etc.) but no corresponding specialty bond funds.[30] Hence, 401(k) participants may rationally avoid equity funds when they face a large number of options if they cannot identify the funds that are suitable for typical investors.

5. Conclusion

Previous research on choice overload has focused exclusively on the possibility that increasing the number of options can reduce participation in a market. Our results establish that the size of the choice sets also impacts what type of options are selected by the market participants. We find that a larger choice set increases the appeal of simple, easy-to-understand, options.

The potential implications of this result are considerable. In many settings, such as Medicare and Social Security, government policy can dramatically increase the number of options that individuals face. Such an increase, unless accompanied with suitable information, could substantially impact the appeal of simpler options even when those are substantially inferior in the long run.

[30] We have looked at whether the presence of more funds increases allocation to index funds relative to actively managed ones, but our estimates were too imprecise for any firm conclusions.

Acknowledgments

We are grateful to Shlomo Benartzi, Ray Fisman, Matthew Gentzkow, Gur Huberman, David Laibson, the editor, the referees, and seminar audiences at Berkeley, Cornell, Wharton, and UCLA for helpful comments and suggestions. John Remarek provided excellent research assistance.

Appendix A

Sample instructions for the extensive condition in 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

Sample instructions the limited condition in 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 falls heads, you receive	If the coin falls tails, you receive
	$5.00	$5.00
	$13.50	$0.00
	$8.75	$3.50

Sample instructions for the extensive condition in 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

Sample instructions for the limited condition in 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
	$1.00	$2.00	$6.75	$7.50	$5.75	$4.75
	$0.00	$0.00	$0.00	$10.00	$10.00	$10.00
	$5.50	$7.50	$0.75	$6.75	$1.00	$6.50

References

Altonji, J., Elder, T., Taber, C., 2005. Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools. Journal of Political Economy 113 (1), 151–184.

Ameriks J. and Zeldes S. P. (2004), "How do household portfolio shares vary with age?" Working Paper.

Benartzi, S., Thaler, R., 2001. Naïve diversification strategies in defined contribution saving plans. American Economic Review 91, 79–98.

Benartzi, S., Thaler, R., 2002. How much is investor autonomy worth? Journal of Finance 57, 1593–1616.

Benartzi, S., Thaler, R., 2007. Heuristics and biases in retirement savings behavior. Journal of Economic Perspectives 21 (3), 81–104.

Benartzi, S., Peleg, E., and Thaler, R. (forthcoming), "Choice architecture and retirement savings plans", in Shafir, Eldar (Ed), The Behavioral Foundations of Policy, Russell Sage Foundation and Princeton University Press.

Bertrand, M., Karlin, D., Mullainathan, S., Shafir, E., Zinman, J., 2010. What's advertising content worth? Evidence from a consumer credit marketing field experiment. Quarterly Journal of Economics 125 (1), 263–306.

Beshears J., Choi, J., Laibson, D., and Madrian, B. (2006), "Simplification and savings", NBER Working Paper 12659.

Boatwright, P., Nunes, J.C., 2001. Reducing assortment: an attribute-based approach. Journal of Marketing 65 (3), 50–63.

Brown, J. R., Liang, N., and Weisbenner S. (2007), "Individual account investment options and portfolio choice: behavioral lessons from 401(k) plans", NBER Working Paper, n. 13169.

Calvet, L.E., Campbell, J.Y., Sodini, P., 2007. Down or out: assessing the welfare costs of household investment mistakes. Journal of Political Economy 115 (5), 707–747.

Calvet, L.E., Campbell, J.Y., Sodini, P., 2009. Fight or flight? Portfolio rebalancing by individual investors. Quarterly Journal of Economics 124 (1), 301–348.

Carroll, G., Choi, J., Laibson, D., Madrian, B., Metrick, A., 2009. Optimal defaults and active decisions. Quarterly Journal of Economics 124 (4), 1639–1676.

Dean, M. (2008), "Status quo bias in large and small choice sets", Working Paper.

Holden, S., VanDerhei, J., Alonso, L., 2009. 401(k) Plan asset allocation, account balances, and loan activity in 2008. ICI Research Perspective 15 (2), 1–62.

Huberman, G., 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., Hepburn, C., 2007. Regret theory and the tyranny of choice. Economic Record 83 (261), 191–203.

Iyengar, S.S., 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., Jiang, W., 2004. How much choice is too much: determinants of individual contributions in 401K retirement plans. In: Mitchell, O.S., Utkus, S. (Eds.), Pension Design and Structure: New Lessons from Behavioral Finance. Oxford University Press, Oxford, pp. 83–95.

Kamenica, E., 2008. Contextual inference in markets: on the informational content of product lines. American Economic Review 98 (5), 2127–2149.

Madrian, S., Shea, D., 2001. The power of suggestion: inertia in 401(k) participation and savings behavior. Quarterly Journal of Economics 116 (4), 1149–1187.

Mottola, G.R., Utkus, S.P., 2003. Can there be too much choice in a retirement savings plan? The Vanguard Center for Retirement Research, Valley Forge.

Sarver, T., 2008. Anticipating regret: why fewer options may be better. Econometrica 76 (2), 263–305.

Shoven, J., Sialm, C., 2003. Asset location in tax-deferred and conventional savings accounts. Journal of Public Economics 88 (1), 23–38.

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

Tversky, A., Shafir, E., 1992. Choice under conflict: the dynamics of deferred decision. Psychological Science 3 (6), 358–361.

Vissing-Jorgensen, A., 2003. Perspectives on behavioral finance: does "irrationality" disappear with wealth? Evidence from expectations and actions. NBER Macroeconomics Annual.