

NBER WORKING PAPER SERIES

HOW GENERAL ARE RISK PREFERENCES? CHOICES UNDER UNCERTAINTY  
IN DIFFERENT DOMAINS.

Liran Einav  
Amy Finkelstein  
Iuliana Pascu  
Mark R. Cullen

Working Paper 15686  
<http://www.nber.org/papers/w15686>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
January 2010

We are grateful to Felicia Bayer, Brenda Barlek, Chance Cassidy, Fran Filpovits, Frank Patrick, and Mike Williams for innumerable conversations explaining the institutional environment of Alcoa, to Colleen Barry, Susan Busch, Linda Cantley, Deron Galusha, James Hill, Sally Vegso, and especially John Beshears, Brigitte Madrian, and Marty Slade, for providing and explaining the data, to Marika Cabral, Tatyana Deryugina, Sean Klein, and James Wang for outstanding research assistance, and to Levon Barseghyan, David Laibson, Dan Silverman, and Jon Skinner for helpful comments. The data were provided as part of an ongoing service and research agreement between Alcoa, Inc. and Stanford, under which Stanford faculty, in collaboration with faculty and staff at Yale University, perform jointly agreed-upon ongoing and ad-hoc research projects on workers' health, injury, disability and health care, and Mark Cullen serves as Senior Medical Advisor for Alcoa, Inc. We gratefully acknowledge support from the NIA (R01 AG032449), the National Science Foundation grant #SES-0643037 (Einav), the Alfred P. Sloan Foundation (Finkelstein), and the John D. and Catherine T. MacArthur Foundation Network on Socioeconomic Status and Health, and Alcoa, Inc. (Cullen). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2010 by Liran Einav, Amy Finkelstein, Iuliana Pascu, and Mark R. Cullen. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

How general are risk preferences? Choices under uncertainty in different domains.  
Liran Einav, Amy Finkelstein, Iuliana Pascu, and Mark R. Cullen  
NBER Working Paper No. 15686  
January 2010  
JEL No. D14,D81,G11,G22

### **ABSTRACT**

We examine the extent to which an individual's actual insurance and investment choices display a stable ranking in willingness to bear risk, relative to his peers, across different contexts. We do so by examining the same individuals' decisions regarding their 401(k) asset allocations and their choices in five different employer-provided insurance domains, including health and disability insurance. We reject the null that there is no domain-general component of preferences. Among the five insurance domains, the magnitude of the domain-general component of preferences appears substantial; we find for example that one's choices in other insurance domains are substantially more predictive of one's choice in a given insurance domain than either one's detailed demographic characteristics or one's claims experience in that domain. However, we find considerably less predictive power between one's insurance choices and the riskiness of one's 401(k) asset allocations, suggesting that the common element of an individual's preferences may be stronger among domains that are "closer" in context. We also find that the relationship between insurance and investment choices appears considerably larger for employees who may be associated with better "financial sophistication." Overall, we view our findings as largely consistent with an important domain-general component of risk preferences.

Liran Einav  
Stanford University  
Department of Economics  
579 Serra Mall  
Stanford, CA 94305-6072  
and NBER  
leinav@stanford.edu

Amy Finkelstein  
Department of Economics  
MIT E52-274C  
50 Memorial Drive  
Cambridge, MA 02142  
and NBER  
afink@mit.edu

Iuliana Pascu  
Department of Economics  
MIT  
50 Memorial Drive  
Cambridge, MA 02142  
pascui@nber.org

Mark R. Cullen  
Stanford School of Medicine  
251 Campus Drive, MSOB x338  
Stanford, CA 94306  
mrcullen@stanford.edu

# 1 Introduction

Standard models in many fields of economics – most notably macroeconomics, finance, public finance, and labor economics – generally use a canonical model for decisions under uncertainty, in which individuals (or households) have a single, concave utility function over wealth, which gives rise to context-invariant risk preferences. Guided by this assumption, standard practice in these literatures is to use external estimates of risk aversion parameters, drawn from a variety of specific contexts, to calibrate their models. At the other end of the spectrum, there is a large literature in psychology and behavioral economics arguing that context is king, and that there is little, if any, commonality in how the same individual makes decisions across different contexts. Where does reality lie relative to these two extremes? Our aim in this paper is to provide new empirical evidence that informs this issue by using a unique data set on thousands of individuals’ insurance and investment choices.

We examine the actual choices that the same individuals make over financial lotteries in different domains. Specifically, we examine the workplace-based benefit choices that Alcoa employees make concerning their 401(k) asset allocations, their disability insurance, and their insurance choices regarding health, drug, and dental expenditures. Using these data we investigate the stability in ranking across contexts of an individual’s willingness to bear risk relative to his peers. In other words, we investigate how well an individual’s willingness to bear risk (relative to his peers) in one context predicts his willingness to bear risk (relative to his peers) in other contexts.

There are several attractive features of our setting for this purpose. First, all the decisions are solely over the extent of exposure to purely financial risk; this reduces concerns about other possible domain-specific components of preferences, such as an individual’s monetary valuation of health. Second, and relatedly, the nature of the contract options makes the different choices within each domain vertically rankable in terms of risk exposure. As a result, we can use these data to investigate the extent to which an individual’s risk aversion relative to his peers in one domain can inform us about his risk aversion relative to those same peers in other contexts. Third, as we shall see, the risk exposure involved in these choices is non-trivial, so that the decisions we observe are economically meaningful. Finally, many of the domains involve risks of similar magnitudes, making the Rabin critique of expected utility theory (Rabin, 2000) less relevant in our setting.

Much of our focus is on attempting to quantify the empirical importance of any individual-specific, domain-general component of preferences. We give considerable attention to trying to develop relevant and useful benchmarks with which to gauge the question of “empirical importance.” This choice of focus is motivated by two factors. First, neither extreme null hypothesis – that of complete consistency or no consistency in preferences across domains – seems (to us) particularly compelling in practice. Reality almost surely lies in between these two extremes. Second, and perhaps more importantly, as we discuss in more detail below, while it is relatively easy to test the null hypothesis that there is no domain general component to preferences (and we will do so), we argue that it is considerably more challenging (perhaps even impossible) to robustly test the other extreme hypothesis that individuals’ decisions are completely consistent across domains. Tests of

the latter hypothesis would inevitably consist of a joint test of the null hypothesis of domain-general preferences as well as a set of difficult-to-test modeling assumptions.

Our results reject the null hypothesis that there is no domain general component of preferences. More interestingly, in our view, we develop several measures that help us assess the extent of this domain-general component of preferences, and we find it to be quantitatively quite important. For example, we estimate that the correlation coefficients that would arise from stylized, calibrated models of coverage choices under fully domain general preferences are quite comparable in magnitude to the actual correlations obtained between choices in different insurance domains (although considerably higher than the estimated correlations obtained between 401(k) portfolio allocations and insurance choices). In addition, we find that one's choices in other insurance domains have about four times more predictive power for one's choice in a given insurance domain than do a rich set of demographics. However, we do find that the riskiness of one's 401(k) portfolio choice has statistically significant but quantitatively much smaller predictive power for one's insurance choices. Interestingly, we also find that predictive power of one's 401(k) portfolio choice for one's insurance choices is systematically greater for individuals who are older, more experienced with the firm, have higher income, or who appear to be more financially sophisticated decision makers (measured based on external proxies in the data). This suggests that such individuals may fit better the canonical model. Overall, we view our findings as generally supportive of the canonical domain-general model of decision making under uncertainty, although they call into question the external validity of risk preferences obtained from contexts that are too far apart or from certain types of populations.

Our study is not unique in its interest in the relative generality of risk preferences across different contexts. Not surprisingly, given its importance, the stability of risk preferences across domains has received considerable attention in the economics literature.<sup>1</sup> Several studies have addressed the stability of risk preferences by investigating individual responses to hypothetical questions regarding financial lotteries across different types of lotteries and over time (Choi et al., 2007; Andersen et al., 2008; Kimball, Sahn, and Shapiro, 2009). Cutler and Glaeser (2005) used a similar approach to investigate correlations in health-related behaviors, for which they use data on self-reported behaviors such as smoking and drinking rather than answers to hypothetical lotteries. The influential paper by Barsky et al. (1997) has analyzed similar hypothetical questions and also validated the responses to some of these questions by investigating whether they are correlated with self-reported behaviors.<sup>2</sup> A recent important study by Dohmen et al. (2009) is probably the closest to our current study; somewhat similar to Barsky et al. (1997), Dohmen et al. (2009) use a large data set of survey responses to hypothetical financial lottery questions and validate

---

<sup>1</sup>Naturally, there is also an important related literature in psychology. Although we do not cover it in detail, many of its features are quite similar to the economics literature we do cover. See, for example, Slovic (1962, 1972a, 1972b) for earlier reviews of this literature and Weber, Blais, and Betz (2002) for a recent paper. See also Schoemaker (1993), who provides an interesting discussion of the contrasting conceptual frameworks by which economists and psychologists address the issue.

<sup>2</sup>See also Chabris et al. (2008) for a similar exercise that focuses on discount rate (rather than on risk preferences).

these responses using self-reported behaviors of a subset of the respondents. Like us, they find an important component of domain-general risk preferences and conclude that although its absolute explanatory power is small, it performs pretty well when compared to other predictors of risk taking.

Our paper differs from this existing literature in several respects. Perhaps most importantly, a common element of all these papers is that they use lab experiments, individual responses to hypothetical questions, or self-reported behaviors. A possible concern with such measures for assessing the domain-generality of an individual’s risk preferences is that there may be important individual-specific elements that affect the mapping from self reported or elicited preferences to actual preferences, which may appear as domain-general preferences. We therefore view the analysis in this paper – which is based on actual market choices and therefore not subject to this concern – as an important contribution.

We are aware of only one other study of the stability of risk preferences across contexts which uses actual market outcomes. Barseghyan, Prince, and Teitelbaum (2009) have recently used data on three similar deductible choices made in the context of auto and homeowner insurance to estimate an individual’s risk aversion in each domain and to test whether they can reject the null that risk aversion is completely general across domains. Our paper differs from theirs in its scope – we look at a much wider range of different domains – and, relatedly, in its focus and empirical approach. While we focus on the extent to which an individual’s risk aversion relative to his peers in a given context can inform us about his risk aversion relative to those same peers in other contexts, Barseghyan, Prince, and Teitelbaum (2009) focus on testing whether the level of risk aversion displayed in different contexts is completely stable across contexts; they reject the null of fully domain general risk aversion. Their question is a more ambitious one, but relies on commensurately greater modeling assumptions as we discuss in more detail below. We therefore view the two approaches (and their results) as highly complementary.

Another contribution of our paper – which also applies to the papers by Chabris et al. (2008) and Dohmen et al. (2009) – is our attempt to quantify the magnitude of any domain general component of preferences by benchmarking it against reasonable alternatives. Most of the studies we have discussed generally find some common element in risk taking within an individual across decisions (or behaviors), although for the most part they tend to argue – on mostly subjective grounds – that this common element is “small.” One of our findings is that the ostensibly “small”  $R^2$ s that many prior papers have found may not in fact be as small when compared to relevant benchmarks.

The rest of the paper proceeds as follows. Section 2 describes our institutional setting and data. Section 3 presents our estimation strategy, emphasizing the difficulty of constructing robust tests for whether preferences are fully domain general, and motivating our focus on providing metrics for quantifying the magnitude of any detected domain-general preferences. Section 4 presents the results, and the last section concludes.

## 2 Setting and Data

We analyze the employee benefit choices from 2004 for the U.S.-based workers at Alcoa, Inc., a large multinational producer of aluminum and aluminum-related products. In 2004 Alcoa had approximately 45,000 active U.S. employees working at about 300 different plants located in 39 states.

We focus primarily on choices made in 2004 because Alcoa introduced a new set of benefit options in 2004, requiring workers to make new, “active” choices in many of the domains we study. As a result, the problems of inferring preferences from “stale” choices is minimized; this could be particularly concerning if individuals might have made their choices about different benefits at different points in time.

We examine employee choices in six different contexts. These include five insurance coverage decisions (health, prescription drugs, dental, and short-term and long-term disability) and one case of asset allocation of the employee’s 401(k) contributions. All insurance choices are made during the “open enrollment” period in November, and apply to the subsequent calendar year. The 401(k) contributions are made automatically every pay-period according to already chosen investment allocations, which in principle could be adjusted at any given time (although in practice only about one quarter of employees in our sample change the allocation of their contributions during a given year). For each choice we observe the menu of options the employee faces (including prices) and the employee’s choice from the menu. We also observe detailed demographic information on the employees and detailed information on the realization of risk during the coverage period.

Prices for the benefit options vary across employees for two reasons. First, for the health, drug, and dental domains, employees have a choice of coverage tier; that is, whether to cover themselves only, or to include their spouse, their children, or the entire family. Throughout this paper we take the coverage tier as given, assuming that it is primarily driven by family structure; we show below that our results are not sensitive to controlling for coverage tier. There is also important cross-sectional variation in the prices associated with each of the insurance options as well as in employer match rates for 401(k) contributions, which we will control for in our analysis using benefit menu fixed effects.<sup>3</sup>

**Baseline sample** Our baseline sample makes a number of restrictions that bring the original 2004 sample of approximately 45,000 active employees down to just under 13,000 employees. First, we restrict our sample to those who were offered the new benefits in 2004; this includes approximately

---

<sup>3</sup>Specifically, the prices faced by the employee are determined by which section of the company the employee is in. Alcoa has about 40 different sections (“business units”). In 2004, each section’s head could select from among the offered “menus” of benefit prices set by Alcoa headquarters (see Einav, Finkelstein, and Cullen (2010) for a much more detailed description). In our sample, there are 20 different possible benefit menus which we control for in the analysis using benefit menu fixed effects. For health, drug and dental the menus vary in the employee premiums. For short-term and long-term disability they vary in the replacement rate associated with the (fixed) premium, although the *incremental* coverage is almost always the same across menus. In the 401(k) domain employees face one of four different possible employer match rates (0, 50%, 75%, or 100%).

all salaried employees but only about one-half of hourly employees, since the benefits provided to union employees (who are all hourly employees) can only change when the union contract expires (so most union employees experienced the change in benefits only in subsequent years). That brings our sample size down to about 26,000 employees. We further restrict the sample to those for which we observe full data on the options they are offered, the choices made, and (for insurance choices) the ex-post realized risk (claims). This precludes, for example, about 8% of the individuals who chose to opt out from Alcoa-provided health and drug insurance coverage and about 11% of employees who chose HMO coverage.<sup>4</sup> We also drop about 22% of the remaining employees who (because of a choice made by their section manager) are not offered long-term disability insurance, as well as the approximately 20% of employees who do not contribute to their 401(k) account.<sup>5</sup> In some of our robustness analyses we add back some of these excluded individuals.

Our final baseline sample contains 12,752 employees. Panel A of Table 1 provides demographic characteristics for this sample. The sample is almost three quarters male and 85 percent white, with an average age of 44, an average job tenure (within Alcoa) of 13 years, and an average annual salary of \$58,400. Only about one-third of the sample is hourly employees and virtually none are unionized (due to our requirement that they face the new benefit option in 2004). The average number of covered individuals per employee is 2.9. Panel B of Table 1 provides summary statistics on the annual payouts for each of the six domains. We now describe the options in each domain in more detail.

**Description of coverage options** As mentioned, we investigate employee’s choices over six different domains. Table 2 attempts to summarize the key features of each domain, with the options enumerated within each domain (as they are in the Alcoa brochures) from the lowest level of coverage (option 1) to the option that offers the most coverage. Appendix Tables A1 and A2 provide more detailed information on each benefit option.

The first domain is health insurance, where employees can choose from among five PPO options.<sup>6</sup> These options only vary in their financial coverage, and (with the exception of option 1) are vertically

---

<sup>4</sup>As is typical in data sets like ours, we do not observe medical expenditures for employees covered by an HMO or who opted out of employer-provided coverage. It is also difficult to analyze the choice of either of these two options since the prices are not known, nor is it entirely clear how to define the “good” being purchased (or to rank it in terms of risk exposure).

<sup>5</sup>Note that the lowest priced option for dental, short-term disability, and long-term disability is free, so that effectively there is no “opt out” option for these domains.

<sup>6</sup>Employees could also choose an HMO or to opt out from health and drug coverage entirely, but those employees who chose these options are excluded from our baseline sample, for reasons described earlier.

rankable,<sup>7</sup> with the deductible level being the key difference.<sup>8</sup> Option 1 stipulates a high annual deductible of 3,000 dollars (for non-single coverage), while option 5 stipulates no deductible. Slightly over half of the employees choose the safest option (option 5), about one quarter choose the second safest option, and about 17 percent choose the least safe option (option 1).

The second domain covers prescription drug coverage, and employees are offered three options that vary in their cost sharing for branded drugs, from 30 percent to 50 percent cost sharing for retail branded drugs (deductible and coverage of generics are the same across options). Almost two thirds choose the safest option and one-quarter choose the least safe option.

The third domain is dental coverage, which offers two options that primarily vary in their annual maximum benefit, of 1,000 vs. 2,000 dollars. About 70 percent of employees choose the safest option.

The fourth and fifth domains are short-term and long-term disability insurance. Short-term disability insurance covers disability-related lost earnings of durations up to six months, while long-term disability insurance covers (less frequent) longer durations. Employees are given a choice of 3 options for each disability insurance coverage, with the replacement rate varying across options. Unlike the first three domains, the pricing and benefits associated with disability insurance are not given in absolute dollars, but rather are proportional to the employee’s annual wage. Thus, the up-front premiums each employee faces vary based on his or her wage, and the benefits are given as “(wage) replacement rates” that are typically 60% and 50% (for short- and long-term coverage, respectively) for the least coverage option and 100% and 70% (respectively) for the options that offer most coverage. About two thirds of the employees choose the highest replacement rate for each option. In each domain, the remaining employees are roughly equally split between the two lower replacement rate options.

The sixth and final domain is the 401(k) asset allocation. As is common in many firms, Alcoa employees are encouraged to contribute every pay-period to their 401(k) account, with Alcoa matching such contributions up to 6%. In our analysis we abstract from the employees’ decisions as to whether and how much to contribute, but rather focus on how contributing employees choose

---

<sup>7</sup>The exception is the cheapest health insurance option (option 1), which is set up as a Health Reimbursement Account (HRA) in which Alcoa contributes each year \$1,250 in tax free money that the employee can use to fund eligible out-of-pocket health care expenses. Any balance remaining at the end of the year can be rolled over to pay for future out-of-pocket costs (as long as the employee remains enrolled in this plan). At retirement (or severance) remaining balances can be used to pay for Alcoa-sponsored retiree health care plan premiums (or toward elected COBRA coverage). Since the financial tax benefits associated with an HRA vary across individuals (based on their marginal tax rates, their expectation regarding future employment with Alcoa, and so on), this introduces a non-vertical component to the health insurance choice. We verify that all the results we report below remain qualitatively similar when we omit the set of individuals who chose this option, but since this set is quite large our preferred specification and analysis simply ignores the tax benefits associated with the HRA.

<sup>8</sup>While there is additional variation across plans in the out-of-pocket maximum and corresponding coverage details of out-of-network expenditure, individuals rarely (less than one percent) reach this out-of-pocket maximum, and only infrequently (less than five percent) use out-of-network services. The out-of-pocket maximum also allows us to abstract from tail risk, which is covered by all options similarly, up to the very similar out-of-pocket maximum across options.



to allocate their contributions across assets. All employees can allocate their contributions and balances among 13 different funds that are available to them, and in principle are allowed to continuously adjust these allocations (although they infrequently do so; for example, only one-quarter of our sample changes its asset allocation during 2004). The funds vary in their riskiness (see Appendix Table A2). To simplify the analysis, we focus on the employees' decisions as to what fraction of their contributions they allocate to the two risk-free funds during 2004.<sup>9</sup> About two fifths of employees allocate none of these contributions to the risk free funds, and about 17% of employees allocate all of their contributions to the risk free funds.

Although describing the options and outcomes in each domain is useful, our understanding of the choices is perhaps best guided by the incremental trade-offs associated with each choice. Columns (2) through (4) of Table 2 provide two (rough) attempts to quantify the relative risk exposure associated with the different choices within a domain. Column (2) does this by reporting the average incremental premium saving in the sample from choosing a given option relative to the least risk exposure option. Columns (3) and (4) report, respectively, the expected and standard deviation of the incremental costs that the employee would face (counterfactually for most of the sample) with the option shown relative to the safest option, if he were to be randomly drawn from our baseline sample. These incremental costs are calculated based on the coverage details and the distribution of realized claims.<sup>10</sup> The most interesting point we take away from Table 2 is that the incremental decisions across each domain are quite comparable in their magnitude, with additional premiums (and associated benefits) ranging from several hundred to a few thousand dollars, annually.

**Attractions of our setting.** The data and setting offer several key attractive features for investigating the extent to which individuals display a common ranking in their risk aversion relative to their peers across domains. First, within all domains, the differences across different choices are purely in the amount of financial risk exposure. They do not involve, for example, differences in access restrictions to health care providers or different service quality by asset fund managers. Such differences would have introduced additional domain-specific elements of the choices that would make interpretation of the results more difficult. Relatedly, since the choices within a domain differ only in the amount of financial risk exposure, they can each be collapsed to a unidimensional vertical ranking of the amount of financial risk one is exposed to in different choices. This makes

---

<sup>9</sup>These two funds are not totally risk free, but they are marketed to employees as the least risky funds, and the standard deviation of their (monthly) returns (0.02 and 0.83) is much smaller than that of the other investment options (which range from 1.36 to 6.71). The results remain similar if we define only the fund with the lowest standard deviation as the risk free allocation, which is not surprising given that the lowest standard deviation fund receives 25% of 401(k) asset allocations, compared to only 4% for the second lowest standard deviation fund. See Appendix Table A2 for more detail.

<sup>10</sup>In our data, expected incremental costs (column (3)) are sometimes higher than incremental premiums (column (2)) suggesting (contrary to fact) that all weakly risk averse individuals will buy the safest option. This is at least partially due to our (unrealistic) simplifying assumption (for the construction of this table) that all individuals are drawn from the same risk distribution.

it relatively straightforward to assess how much more likely it is for individuals who assume more vs. less risk compared to their peers in one domain to assume more vs. less risk in another domain compared to their peers.

Second, as shown in columns (3) and (4) of Table 2, all of the domains are plausibly valuable and sensible insurance from an economic standpoint. That is, they all represent potentially large expenditures with real ex ante uncertainty to the individual. For example, the coefficient of variation of incremental costs (computed based on columns (3) and (4)) is always greater than one third, and mostly greater than one. This is a much more appealing setting for studying the extent to which choices across domains display a common risk aversion component than looking at settings in which it is unclear why individuals are buying insurance in the first instance, such as insurance for internal wiring protection (as in, e.g., Cicchetti and Dubin, 1994) and other types of “insurance” products that cover against very small losses, which Rabin and Thaler (2001) argue is where people are perhaps most likely to depart from expected utility theory.

Third, as discussed earlier (and shown in Table 2, columns (2) and (3)), the choices within a domain are over similarly sized risks, making our analysis less vulnerable to Rabin’s critique of expected utility theory (Rabin, 2000).

Fourth, many of the benefit options are entirely new in 2004, which means that for these benefit options we are looking at decisions made all at the same time period and do not have to worry about “stale” decisions in some domains reflecting a combination of inertia and outdated risk preferences. Specifically, the health, drug and dental options were all completely new, while the disability options remained the same but their prices changed; the 401(k) options did not change.<sup>11</sup> As a further check against the possibility of “stale” decisions (particularly for 401(k) allocations and potentially disability choices), we show in our robustness analysis that the main results look similar when restricted to a sample of new hires, for whom decisions in all six domains had to be made recently.

Fifth, and relatedly, with the exception of the 401(k) asset allocation decisions, the nature of the employee benefit selections eliminates many potential domain-specific elements of the choice; all the insurance benefits are presented in the same format (all on the same benefit worksheet) and must be chosen during the same open enrollment period. Thus, we do not have to worry for example about time-varying events, differential effort or ability of insurance agents, etc.

Sixth, there is interesting variation across the six domains in the “closeness” of the domains. In particular, it seems that some domains (such as short-term and long-term disability insurance) are quite similar while others (such as health insurance and 401(k) decisions) are more different. Therefore it is interesting to see if the extent of correlation in choices within an individual across domains varies by their relative “closeness.”

---

<sup>11</sup>We also know the default options for each domain which are: health insurance option 4, drug insurance option 3 single coverage, and for dental, short- and long-term disability the default is one’s prior year’s choice if he or she was previously employed (or no coverage, lowest option, and middle option respectively if they are a new hire). Of course, people in these allocations may also have chosen them actively. In our robustness analysis we explore sensitivity to excluding people who, based on their allocations, may not be active choosers.

Finally, but very importantly, the data are extremely clean and complete. We observe all the details of the choice set, the choice made, the setting in which the choice is made, a measure of risk occurrence, and quite rich demographic information.

### 3 Empirical strategy

A natural way to evaluate the stability of risk preferences across domains is to write down a model of consumer behavior, use the data and the model to obtain estimates for risk aversion for each consumer in each domain, and then compare the estimates obtained for the same individual in different domains. However, for relatively disparate domains such as ours, such an exercise will almost surely be sensitive to the modeling assumptions. For example, health, dental, and prescription drug insurance are associated with absolute (dollar) costs and distribution of risk, while disability and investment decisions are driven by risks that are proportional to the employee’s wage and wealth, respectively. More generally, estimating the distribution of risk aversion from individuals’ insurance choices requires a domain-specific model of ex-ante heterogeneity in risk (Cohen and Einav, 2007; Barseghyan, Prince, and Teitelbaum, 2009).

Our strategy is therefore to avoid most of these modeling assumptions by focusing on the within-person correlation in the ordinal ranking of the riskiness of the choice an individual makes across different domains. In other words, we ask whether individuals who appear to be more willing to bear risk than their peers in one context are also more willing to bear risk in another context. Admittedly, this is a more modest test of the stability of risk preferences. For example, we could find that individuals are stable in their relative ranking, although the entire distribution of willingness to assume risk is dramatically different across domains (but rank preserving). The advantage of this approach, however, is that it allows us to make inferences that are much more robust to various difficult-to-test modeling assumptions. In particular, the approach “only” requires that any unobserved individual- and domain-specific components of demand in a given domain are rank preserving; it does not require us, for example, to take a stand on the nature of the utility function (e.g. constant absolute vs. constant relative risk aversion) or a model by which we can map the distribution of ex-post claims experience into ex-ante private information about risk type.

**Quantifying vs. testing** Our primary interest is in developing reasonable benchmarks against which one could compare and contrast our estimates of the correlation in the ordinal ranking of the riskiness of one’s choices across domains. For reasons both of interest and practicality, we focus on ways of quantifying and interpreting the magnitude of our estimates rather than testing the two possible extreme nulls. At one possible extreme of our analysis (which for lack of a better term we would call the “neoclassical” model), the rank correlations would be one and (relative) risk preferences would be “fully” domain general. In the other extreme (which for lack of a better term we would call the “narrow framing” model) we assume that there is no domain-general component to preferences, so one’s risk aversion in one domain has no predictive power for how one makes choices in another domain, and the correlation would be zero. As noted earlier, we find neither of

these extreme hypotheses particularly compelling as representations of reality, which serves as part of our motivation to focus on quantifying the extent of domain generality rather than testing either extreme null.

Relatedly, a robust test of the null of full domain generality is elusive, since even if preferences are fully domain general, the rank correlation need not be one. For example, even if risk preferences are fully domain general, any discreteness and non-linearity in the function that maps risk aversion to choices would make the correlation estimates lower, potentially by a substantial amount. To illustrate this with a concrete example, suppose we observe  $N$  individuals making choices in two domains ( $j$  and  $k$ ), each of which offers two discrete choices, with choice 1 exposing the individual to more risk than choice 2. Even if preferences are fully domain general, it is possible that due to the different pricing of options in the two different domains, in domain  $j$  the lowest risk aversion individual chooses option 1 while all  $N - 1$  other individuals choose option 2, while in domain  $k$  the highest risk aversion individual chooses option 2 and all  $N - 1$  other individuals choose option 1. While this allocation is consistent with the underlying model of fully domain general preferences, the correlation of choices across the two domains will approach zero as  $N$  gets sufficiently large.

In addition, in a fully domain general model with a single utility function over wealth, insurance decisions are inter-related, and one essentially chooses a portfolio of insurance positions. In other words, risk exposure in one domain may affect (with ex ante ambiguous sign) one’s willingness to bear risk in another (even independent) domain (e.g., Gollier and Pratt, 1996; Guiso, Jappelli, and Terlizzese, 1996). This so-called “background risk” problem introduces yet another reason why fully domain general preferences need not produce a rank correlation of one across domains.

Thus, testing for a correlation of one – or comparing the estimated correlations against a fully domain general benchmark of one – is not, we would argue, meaningful or appropriate. In the next section we offer several alternative possible benchmarks that can help inform the extent of domain generality. By contrast, rejecting the null of a zero correlation is a more meaningful rejection of the null of no domain-general component to preferences. Of course, it is subject to the concern that non-preference factors may introduce correlations across domains. In the case of insurance a natural suspect is potential correlation in risks across domains, which is why we consider the correlation between 401(k) asset allocations and any given insurance choice as the cleanest test of this null.

**Econometric specification** Given our interest in the extent to which individuals’ ranking in their risk aversion relative to their peers displays a common component across domains, a natural empirical approach is to examine the rank correlation within individuals across domains in their choice from among the (vertically ranked) options in a domain. We therefore begin by reporting pairwise Spearman rank correlations across domains. A disadvantage to this approach, however, is that it does not readily lend itself to controlling for potentially important covariates nor does it lend itself as easily to a construction of comparative benchmarks with which to gauge the relative importance of the domain general component of risk preferences that we detect.

We therefore also examine the correlation structure of the error terms in a multivariate regression

of the form:

$$\begin{bmatrix} choice_i^{Health} \\ choice_i^{Drug} \\ choice_i^{Dental} \\ choice_i^{STD} \\ choice_i^{LTD} \\ choice_i^{401(k)} \end{bmatrix} = \begin{bmatrix} \beta^{Health} \\ \beta^{Drug} \\ \beta^{Dental} \\ \beta^{STD} \\ \beta^{LTD} \\ \beta^{401(k)} \end{bmatrix} \cdot x_i + \begin{bmatrix} \varepsilon_i^{Health} \\ \varepsilon_i^{Drug} \\ \varepsilon_i^{Dental} \\ \varepsilon_i^{STD} \\ \varepsilon_i^{LTD} \\ \varepsilon_i^{401(k)} \end{bmatrix} \quad (1)$$

where  $x_i$  is a vector of control variables (which is the same in all regressions),  $\beta$  is a vector of domain-specific coefficients, and the main object of interest is the correlation matrix of the residuals. In practice, the rank correlations and the correlations from the multivariate regression produce extremely similar results.

In our baseline specification we enumerate the choices from 1 to  $n$  in each domain (as in Table 2), and estimate equation (1) using (multivariate) least squares. This specification has the disadvantage that it treats an ordinal choice as a cardinal variable. However, since the results are remarkably similar to those of the Spearman rank correlation (as well as to a system of ordered probit equations on which we report in the robustness section), we are comfortable using this simpler approach to explore sensitivity to the choice of control variables ( $x_i$ ) and to provide benchmarks for gauging the magnitude of the correlations we estimate.

In our baseline multivariate regression specification we include controls (in the form of dummies) for the menu of benefits the employee faced (described above). We also explore the sensitivity of our results to the inclusion of controls that proxy for individual risk in a given domain. Because standard theory models insurance choices as driven by risk and risk aversion, isolating risk aversion requires controlling for risk. In doing so, we control for both predictable risk as well as for idiosyncratic risk. For the latter we use the realization of risks in the subsequent coverage period. For the former, we use our data to run a predictive statistical model of realized risks on a flexible functional form of our observables, and then generate the model predictions and use these predictions as controls.<sup>12</sup> Interestingly, controlling for risks makes little difference, so we are not particularly worried about how much better we might have been able to predict risk with even more flexible specifications.

## 4 Results

### 4.1 Baseline correlations

Table 3 presents the main correlation results. Panel A shows the full set of Spearman rank correlation coefficients between each pair of domains. It also reports (at the bottom) the simple average of the fifteen correlation, as a single summary measure. Panel B shows the correlations from the baseline multivariate regression described above in equation (1). The results are remarkably similar

---

<sup>12</sup>Specifically, risk is predicted by regressing realized risk on cubic splines in age, wage, and job tenure, dummy variables for gender, race, employee type (hourly or salary), union status, single coverage for health benefits, family size, state fixed effects, and interaction between age and these dummy variables.

across the two panels which is why (as discussed above) we use the multivariate regression in Panel B as our baseline specification.

In all of the pairs, we can reject the null hypothesis of a correlation of 0. That is, we can reject the null of no domain general component of choice. Viewed alternatively, we find that one's coverage choice in every other domain has some predictive power for his or her choice in a given domain.

Of course, an important caveat to this conclusion is that correlations in choices across insurance domains could reflect correlation in underlying (unpriced) risk across the insurance domains; such an issue does not arise in the context of the correlation between 401(k) portfolio allocation and choices in an insurance domain, making this perhaps the most compelling context to test the null of complete domain specificity. To try to address the concern about underlying risk correlations across insurance domains, Panel C of Table 3 reports the results from a variant of the baseline multivariate regression in which we control (as explained earlier) for both predicted and realized risk in *all* domains in *each* equation. The magnitude of the correlations generally remains almost the same, with only a slight decline (the decline is to be expected, given that these risks are positively correlated). While predicted and realized risk do not control perfectly for one's ex ante risk expectations, the small effect that these controls have on the correlation pattern suggests that these correlations are more likely to capture correlation in underlying risk preferences. This is also consistent with recent results – in the context of a fully specified model – that heterogeneity in risk preferences plays a much greater role than heterogeneity in risks in explaining the heterogeneity in insurance coverage choices (Cohen and Einav, 2007; Barseghyan, Prince, and Teitelbaum, 2009).

Although the finding that risk preferences are correlated across domains may be viewed as hardly surprising, we are not aware of any other work that documented it, at least in the context of actual choices (that is, non-experimental and non-survey), across a broad range of contexts. We believe this finding provides one of the most compelling evidence against the null of no domain-general component to choice.

In all three panels of Table 3 we see that the average pairwise correlation is 0.16 to 0.19. Not surprisingly, there is a pronounced pattern of substantially higher correlation coefficients between pairs that are more “similar.” For example, in panel B (which we will use as our baseline specification), the correlation between drug and health coverage choices is 0.45 and the correlation between long-term and short-term disability insurance choices is 0.60. By contrast, medical insurance and short-term disability insurance show only a 0.19 correlation and the lowest pair-wise correlations are between the share of risk free assets in one's 401(k) portfolio and any of the insurance coverage choices (all of which are below 0.06). Of course, it is not clear how informative this finding is since comparisons of correlations between different pairs are difficult to interpret due, for example, to differences in the discreteness and pricing of the relative options in each domain.

## 4.2 Benchmarks to help interpretation

While we interpret the results presented in Table 3 as supporting the claim that there exists a domain general component of risk preferences, as noted at the outset, our primary interest is in developing useful metrics for gauging the quantitative importance of domain-general element of preferences that we have detected. As we emphasized in Section 3, comparing the estimated correlations to the benchmark of a correlation of 1 does not provide a meaningful assessment of the extent of domain generality of preferences, as the correlation in choices may be substantially less than 1 even if preferences are fully domain general.<sup>13</sup> In this section we therefore suggest three potentially useful metrics for gauging the extent of domain general preferences implied by the correlation estimates in Table 3. Our general conclusion from these benchmarks is that, contrary to our expectations, the reported average correlations of 0.16-0.19 are in fact quite high, and suggestive of an important domain-general component of risk preferences.

**Simulated correlations from a calibrated model.** One natural way of gauging the magnitude of our estimated correlation coefficients is to compare them to the correlation coefficients that would arise from a calibrated model of coverage choice. The objective of the calibration exercise is to obtain some benchmarks for the correlation coefficients between choices generated by a model with completely domain-general risk preferences, but subject to the non-linearities and discreteness transformations that arise because of the structure of the insurance options, and the decision process.

We use the short-term and long-term disability insurance settings and simulate optimal coverage choices from among three options similar to those described in Table 2. We choose these two domains because their similarity allows us to use the same modeling assumptions in both. We assume an expected utility model with constant relative risk aversion (CRRA), we calibrate the distribution of risk aversion, and use the claims data to define eight risk groups based on demographics. We then simulate coverage choices from the modal menu in both these domains and calculate the correlation in coverage choices using different assumptions about the correlation in risk. We also compute correlations between the short term disability insurance coverage choice and the underlying coefficient of risk aversion; this correlation is conceptually similar to the correlation between the discrete insurance coverage choices in our data and the continuous measure we use to capture risk aversion in the 401(k) setting. The appendix describes these exercises in more detail.

The correlations we obtain in the first exercise, across short term disability insurance and long term disability insurance vary between 0.18 (when we assume that the individual risk group is independent across the two choices) and 0.55 (when we assume the risk groups are the same). When

---

<sup>13</sup>Relatedly, one might think that a natural way to gauge the extent of domain general preferences is via factor analysis. Greater domain generality would imply greater importance of the first factor relative to the incremental importance of additional factors. We performed a principal components analysis and found that the first factor explains 35 percent of the observed variation in the six domains, compared to only 20 percent for the next most important factor. We note, however, that such analysis raises the same difficulties for interpretation; without proper benchmarks it is hard to evaluate whether 35 percent is high or low.

we control for risk (using risk group fixed effects) the latter correlation drops to 0.47. One could think of comparing these numbers to the correlation we report between short-term and long-term disability coverage of around 0.6. We have also experimented with correlating simulated choices from two different menus of the same coverage (of short-term disability insurance); this exercise allowed us to assume identical risks across the choices (as the domain is the same), and even in such cases the correlation in choices remained below 0.79, and dropped to 0.50 after controlling for risk group fixed effects. We view this qualitative analysis as suggesting that a correlation of 0.6 is quite high.<sup>14</sup>

Finally, the correlation we obtain between the choice from the discrete short term disability insurance menu and the underlying risk aversion coefficient is 0.14. This may be compared to the much smaller correlations we empirically find between the discrete insurance choices and the continuous 401(k) asset allocation. This is consistent with our earlier interpretation that the correlation between 401(k) and the insurance choices is indeed quite low if one has in mind a model with fully domain-general risk preferences.

**Predictive power of covariates.** Another way of gauging the importance of the domain general element of preferences is by comparing the predictive power of choices in other domains to the predictive power of demographic covariates. Table 4 reports these results. For each domain, it reports the adjusted  $R^2$  from a multivariate regression of the (ordinal) coverage choice in this domain on different subsets of covariates. All regressions are done on the residual coverage choice (after partialing out the menu fixed effects). As one can see, the explanatory power (measured by the adjusted  $R^2$ ) of the choices in other domains (row 1) is much greater for predicting one’s insurance choice in a different domain than the predictive power of one’s risk type (row 2), or one’s detailed demographics (row 3). For example, the predictive power of choices in other domains is at least four times greater than the predictive power of demographics in predicting the choice in a given insurance domain. Even when we limit the choices in other domains to exclude the most related coverage choice (row 4), the predictive power of the remaining choices is at least 1.5 times higher than that of demographics for the choice in a given insurance domain.

The case of 401(k) is a noted exception to this pattern. The explanatory power of the insurance choices (row 1) is an order of magnitude lower than that of demographics. This is not a particularly surprising pattern, given the relative “distance” between 401(k) and all the other choices, as well as potential differences in the timing (or framing) of the decision, and potential age-based preferences for the (longer horizon) 401(k) investments.

**Within person correlation for the same domain over time.** Yet another natural benchmark for gauging the extent of domain general preferences is to compare the correlation within person in choices across domains at a point in time to the correlation within person in choices in a given

---

<sup>14</sup>This exercise was also reassuring that – despite the discrete choice set – the null of no domain general preferences manifests itself to essentially zero correlation when we assume independent risk aversion (and independent risk groups) across domains.



domain over time. Here again we can take advantage of the new benefit design that Alcoa introduced in 2004, and compute the correlation for health insurance choices between 2003 and 2004. In the “old” benefit design (of 2003), individuals could choose from among three different coverage options (compared to five in the new design), with variation in out-of-pocket maximum being binding and important. These three options were also vertically rankable from least to most coverage, just like other domains in 2004, thus providing a similar structure, and a comparable benchmark. The correlation we find between health insurance choices (of the same employee) in 2003 and 2004 is 0.198. This is similar to (or smaller than) the across-insurance-domain correlations reported in Table 3 which range from 0.16 to 0.60.<sup>15</sup>

### 4.3 Correlations for different groups

In addition to exploring the extent of domain general preferences for the entire sample, we also examined how this varies across different identifiable groups. We report the main results in Table 5, which shows selected correlations for different pairs of groups of employees. While many pairwise correlations seem to be quite similar across groups, the most striking pattern in Table 5 is in column (5), which shows the pattern across groups in the correlation between health insurance coverage choices and 401(k) allocation decisions. A similar pattern is observed across these groups in the correlations between other insurance choices and the 401(k) decisions (not shown in the table in the interest of space).

Quite remarkably, the results in column (5) reveal a consistent pattern that individuals whom one might ex ante classify as likely to make better financial decisions tend to have noticeably higher correlations between health insurance choices and 401(k) decisions. This is true for older individuals relative to younger individuals, for individuals with longer tenure with Alcoa (who perhaps understand the “system” better), individuals with higher wages, and individuals who tend to avoid what economists often view as unsophisticated financial behavior, such as not rebalancing the portfolio regularly and allocating some of their contributions to Alcoa stock.

One way to interpret these findings is that while the correlation between insurance and 401(k) investment choices is low in the overall sample, we find a greater degree of domain-general risk aversion once we focus on individuals who exhibit more “financial literacy,” or at least seem to pay more attention to their investment decisions. An alternative, plausible interpretation is that these results suggest less error in the mapping from “true” underlying risk preferences to choices for individuals who appear to be more “financially literate.” This latter interpretation is consistent with a growing body of empirical work suggesting that the propensity to succumb to psychological biases or to make mistakes in financial planning is higher for individuals of lower cognitive ability (Benjamin, Brown, and Shapiro 2006) and for individuals of lower financial literacy or planning propensity (Ameriks, Caplin, and Leahy 2003; Lusardi and Mitchell, 2007). Either interpretation

---

<sup>15</sup>One could also investigate correlation in choices over time without any change in benefit design. The concern about such an exercise is that employees don’t make “active” choices in subsequent years, which is precisely the reason that made us use the new benefit design for the baseline exercise. Indeed, when we examine such correlations (looking at years 2004 and 2005), we obtain correlation coefficients of 0.85-0.9, presumably due to inertia.

suggests that one might want to exercise more caution in using specific revealed preference estimates to calibrate risk aversion levels in economic models, when they are applied to less sophisticated populations.

Finally, in a different vein, we note a reassuring result in Table 5 that there is no evidence of higher correlations for the newly hired than for more tenured employees, either in general or particularly for the correlation between health insurance choices and 401(k) decisions. This is reassuring since it alleviates concerns that health insurance might reflect a “new” or “active” choice while the 401(k) asset allocation decisions might reflect inertial behavior.

#### 4.4 Robustness

We explored the robustness of our main correlation results (in Table 3) to various alternative specifications and samples. Table 6 summarizes the results of these analyses. As in Table 5, in the interest of space, we do not report every pairwise correlation, but instead report the average correlation and the correlations of three representative (and interesting) pairs. We explore two main types of sensitivity analysis: alternative specifications and alternative samples. Unless otherwise specified, each row represents a single variant from the baseline specification. Overall, the results suggest relatively little sensitivity to these choices.

Row 1 shows our baseline specification from Panel B of Table 3. This reports the correlation of the residuals from the multivariate regression given in equation (1); the control variables in this baseline specification are the benefit menu fixed effects. We previously showed that the results were robust to using the Spearman rank correlation without any control variables (Panel A of Table 3) or to adding additional control variables for risk type to the baseline specification (Panel C of Table 3). Panel A of Table 6 further investigates the sensitivity of the results to alternative specifications. Row 2 reports the correlations of residuals from estimating a mixed system of ordered probits for the five insurance domains (rather than linear regression models in the baseline specification) and a linear model for the 401(k) domain. This more appropriately reflects the ordinal nature of the dependent variables in the insurance domains.<sup>16</sup> Row 3 reports results from a variant of our baseline specification in which we discretize the 401(k) asset allocation decision and turn it into an ordinal measure, so as to investigate the extent to which the relatively low correlation between 401(k) and insurance domains is affected by the continuous, cardinal nature of the 401(k) domain. Specifically, we take the continuous measure of the percentage of employee contributions allocated to the safe funds, and convert it to a discrete integer between 1 to 5, with each corresponding to a 20% bin from riskiest to safest (1 corresponds to 0-20% in risk free assets, 2 corresponds to 20-40%, and so on). In row 4 we investigate the sensitivity of our results to including indicator variables for the (four) coverage tiers (single coverage, employee plus spouse, employee plus children, and

---

<sup>16</sup>We estimate this mixed multi-equation model using maximum likelihood. The estimation is performed using the `cmp` user-provided package in STATA (see: <http://ideas.repec.org/c/boc/bocode/s456882.html>). All the regressions include pricing menu fixed effects, and we exclude all the menus that were offered to fewer than 100 people. This reduces the sample size by 86 employees.

family coverage). Finally, to investigate concerns about whether our benefit menu fixed effects fully capture differences in choices due to prices, in row 5 we limit the sample to those who faced the prices in the single largest menu (about 60 percent of our baseline sample) and re-estimate our baseline specification (without the menu fixed effects) on this subsample.

Panel B of Table 6 explores the sensitivity of our baseline specification to alternative sample definitions. In rows 6 through 8 we add back in various employees who were excluded from the baseline sample. In row 6 we include those employees who opted out of the health insurance and drug insurance plans, or who chose an HMO for these plans. In row 7 we include employees who did not contribute to their 401(k) plan in 2004, and in row 8 we include those employees who were not offered long-term disability insurance. In each case, we omit from the analysis the affected domains (health and drug in row 6, 401(k) in row 7, and long-term disability in row 8). As a result, comparison of the average correlation to that in the baseline may be misleading, but the pairwise ones are still informative, and we also report the comparable average correlation in the baseline specification. In row 9 we exclude from our analysis individuals who chose health insurance option 1, the lowest coverage option, which, because it is bundled with a Health Retirement Account component, is not fully vertically rankable (see Appendix Table A1 and discussion in text above). In row 10 we limit the sample to the slightly under 10 percent of the sample who were new hires in 2004. As discussed earlier, a primary motivation for this analysis is to see if 401(k) contribution allocations are more highly correlated with insurance choices when the 401(k) choice (like the insurance choice) must be a new and “active” decision. In practice, there is no evidence that differences in timing of the decision is driving down the correlation between 401(k) asset allocation and insurance coverage. Finally, in row 11 we exclude the roughly 11 percent of individuals who might have been “passive” choosers, given that all their coverage decisions in the insurance domains were consistent with the default options.

## 5 Conclusion

This paper has investigated the extent to which individuals display a stable ranking in their risk preferences relative to their peers in making choices over different financial lotteries. In contrast to much of the prior work on this topic we have focused on real-world choices as opposed to hypothetical questions, laboratory experiments, or self-reported behaviors. We have also chosen to follow an empirical strategy which investigates to what extent individuals’ willingness to take risks relative to their peers is stable across domains, rather than to attempt to test the null of completely domain general preferences. While our chosen approach sets a lower hurdle for a canonical domain-general risk aversion theory to pass, it also relies more on the data and less on modeling assumptions, making it more robust. An important portion of the paper has tried to develop useful benchmarks which would allow us to gauge the magnitude of any domain-general component of preferences.

We performed our analysis in the context of the workplace-based benefit choices of Alcoa employees regarding their 401(k) asset allocations, their short- and long-term disability insurance, and their health, drug, and dental insurance choices. This setting has the attraction that the decisions

are all over purely financial risk, the choices within each domain are easily vertically rankable in terms of risk exposure, and the domains involve risks of similar (and quantitatively non trivial) magnitudes. In this setting, we reject the null hypothesis that there is no domain general component to preferences and, more interestingly, we find that the extent of the domain-general component appears to be substantively important. For example, we find that the correlations we document are similar in magnitude to those that would occur from simulated choices in a calibrated model that subjects individuals to the same options we observe in the data; we also find that one’s choices in other insurance domains have about four times more predictive power for one’s choice in a given insurance domain than do a rich set of demographic variables.

On the other hand, we also find evidence of non-trivial context specificity. In particular, we find that the riskiness of one’s 401(k) asset allocation decisions has considerably less predictive power for one’s insurance choices than do other insurance choices (or demographics). More generally, even within the insurance domains we find a higher correlation in choices that are “closer” in context (such as health insurance and drug insurance, or short-term and long-term disability insurance) than ones that are further apart (such as health insurance and disability insurance).

An implication of these findings is that when calibrating models of economic behavior – be it insurance demand, savings, labor supply, etc. – it may be important to try to calibrate the model using preference estimates taken from similar contexts. Calibration work is ubiquitous in the fields of insurance, public finance, and macroeconomics. To our knowledge, the vast majority of this work (including our own past work) attempts to calibrate the models using “consensus” parameter estimates (or ranges of estimates) from “the literature at large” rather than estimates from more similar contexts.

Relatedly, our findings suggest that the extent of domain-generality of preferences may be systematically higher for certain types of individuals – including wealthier individuals, older individuals, more financially sophisticated individuals – than others. This suggests that the use of a domain-general preference parameter in calibration work may be a more reasonable assumption for certain types of populations. The question of what contexts and for what types of individuals domain-specificity is largest remains an interesting question for further exploration.

## 6 References

Ameriks, John, Andrew Caplin and John Leahy (2003). “Wealth Accumulation and the Propensity to Plan.” *Quarterly Journal of Economics* 68:1007-1047.

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rusterom (2008). “Lost in space: are preferences stable?” *International Economic Review*, 49(3), 1091-1112.

Barsky, Robert, Thomas Juster, Miles Kimball, and Matthew Shapiro (1997). “Preference parameters and behavioral heterogeneity: an experimental approach in the health and retirement study.” *Quarterly Journal of Economics*, 112(2), 537-579.

Barseghyan, Levon, Jeffrey Prince, and Joshua Teitelbaum (2009). “Are risk preferences stable across contexts? evidence from insurance data.” Mimeo, Cornell University.

Benjamin, Daniel, Sebastian Brown, and Jesse Shapiro (2006). “Who is ‘Behavioral’? Cognitive Ability and Anomalous Preferences.” Mimeo, Harvard University.

Chabris, Christopher, David Laibson, Carrie Morrie, Jonathan Schuldt, and Dmitry Taubinsky (2008). “Individual laboratory-measured discount rates predict field behavior.” *Journal of Risk and Uncertainty*, 37(2), 237-269.

Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv (2007). “Consistency and heterogeneity of individual behavior under uncertainty.” *American Economic Review*, 97(5), 1921-1938.

Cicchetti, Charles J., and Jeffrey A. Dubin (1994). “A microeconomic analysis of risk aversion and the decision to self-insure.” *Journal of Political Economy*, 102(1), 169–186.

Cohen, Alma, and Liran Einav (2007). “Estimating risk preferences from deductible choice.” *American Economic Review*, 97(3), 745-788.

Cutler, David, and Edward Glaeser (2005). “What explains differences in smoking, drinking, and other health related behaviors?,” *American Economic Review papers and proceedings*, 95(2), 238-242.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jurgen Schupp, Gert G. Wagner (2009). “Individual risk attitudes: measurement, determinants and behavioral consequences.” *Journal of European Economic Association*, forthcoming.

Einav, Liran, Amy Finkelstein, and Mark Cullen (2010). “Estimating welfare in insurance markets using variation in prices.” *Quarterly Journal of Economics*, 125(3), forthcoming.

Finkelstein, Amy, and Kathleen McGarry (2006). “Multiple dimensions of private information: evidence from the long-term care insurance market.” *American Economic Review*, 96, 938-958.

Gan, Li, Michael Hurd, and Daniel McFadden (2005). “Individual subjective survival curves.” in *Analyses of the Economics of Aging*, David Wise (ed.). Chicago, IL: University of Chicago Press.

Guiso Luigi, Tullio Jappelli, and Daniele Terlizzese (1996). “Income Risk, Borrowing Constraints, and Portfolio Choice.” *American Economic Review* 86(1): 158-172.

Gollier, Christian, and John W. Pratt (1996). “Risk Vulnerability and the Tempering Effect of Background Risk.” *Econometrica* 64(5): 1109-1123.

Hurd, Michael D., and Kathleen McGarry (2002). “The predictive validity of subjective probabilities of survival.” *Economic Journal*, 112(482), 966-985.

Kimball, Miles S., Claudia R. Sahm, and Matthew D. Shapiro (2009). “Risk preferences in the PSID: individual imputation and family covariation.” NBER Working Paper No. 14754.

Lusardi, Annamaria, and Olivia S. Mitchel. 2007. “Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth.” *Journal of Monetary Economics* 54: 205 - 224.

Rabin, Matthew (2000). “Risk aversion and expected-utility theory: a calibration theorem.” *Econometrica*, 68(5): 1281–1292.

Rabin, Matthew, and Richard H. Thaler (2001). “Anomalies: risk aversion.” *Journal of Economic Perspectives*, 15(1), 219–232.

Schoemaker, Paul J. H. (1993). "Determinants of risk-taking: behavioral and economic view." *Journal of Risk and Uncertainty*, 6, 49-73.

Slovic, Paul (1962). "Convergent validation of risk-taking measures." *Journal of Abnormal and Social Psychology*, 65, 68-71.

Slovic, Paul (1972a). "Information processing, situation specificity, and the generality of risk-taking behavior." *Journal of Personality and Social Psychology*, 22(1), 128-134.

Slovic, Paul (1972b). "Psychological study of human judgemental implications for investment decision making." *Journal of Finance*, 27(4), 779-799.

Smith, V Kerry, Donald Taylor and Frank Sloan (2001). "Longevity expectations and death: can people predict their own demise?" *American Economic Review* September pp.1126-1134.

Weber, Elke U., Ann-Renee Blais, and Nancy E. Betz (2002). "Measuring risk perceptions and risk behaviors." *Journal of Behavioral Decision Making*, 15, 263-290.

## Appendix: Description of the calibration exercise

The objective of the calibration exercise is to obtain a benchmark for the correlation coefficients that would be produced if the data were generated by a model with completely domain-general risk preferences, but were subject to the non-linearities and discreteness transformations that arise from the ordinal coverage choice data. We focus on short- and long-term disability, which are the domains that are most similar to each other in their structure of choices and risks (so we can rely on a single choice model for both domains, rather than on a domain-specific model, which is not ideal). We estimate the correlation in the simulated choices between the modal short-term disability menu (of 60%, 80% and 100% replacement rates) and the modal long-term disability menu (of 50%, 60% and 70% replacement rates), using the observed prices. In the text we also report correlation results between the short-term disability choices and the underlying (continuously distributed) risk aversion parameter, as well as between short-term disability choices from two different menus (the other being 40%, 50%, and 60% replacement rates).

Our calibration exercise assumes a constant relative risk aversion (CRRA) per-period utility function, whereby the expected utility from a given disability insurance contract (which specifies a given wage replacement rate and is associated with a given annual premium) is

$$Eu(c) = E \left[ \left( 1 - \tilde{d} + replacement\_rate * \tilde{d} - premium \right)^{1-\gamma} \right], \quad (2)$$

where  $\tilde{d}$  is the (ex-ante random) fraction of days in a year the employee claims (due to disability), the premium is measured per dollar of (annual) wage, and  $\gamma$  is the coefficient of relative risk aversion. The individual is maximizing expected utility over the duration under consideration, which we assume to be one year for short-term disability and four years for long-term disability (as after about four years, our claim data is truncated, although only few disability claims in the data remain active that long). We assume a discount factor of .95 for long-term disability.

We assume the distribution of  $\gamma$  in the population to be lognormal with parameters  $\mu = -1.2089$  and  $\sigma = 2.1483$ . These parameters were chosen to produce an average relative risk aversion coefficient of 3 (as often used in the literature) and a coefficient of variation of risk aversion of 10, which matches the estimates reported by Cohen and Einav (2007). We note that while Cohen and Einav (2007) mention higher numbers of relative risk aversion, they essentially estimate absolute risk aversion, so mapping it to this lower level of relative risk aversion amounts to simply assuming lower relevant wealth.<sup>17</sup> To calibrate the distribution of risk (missed days), we use the risk realization of short- and long-term disability in the data to define eight risk groups based on demographics (using all combinations of gender, race, and employment status indicators), which produces a fairly large heterogeneity in ex ante risk across individuals. We assume a sample size identical to our baseline sample (12,752) and for each individual in the calibrated sample we draw a risk aversion coefficient from the distribution of  $\gamma$ , assume that he or she knows the distribution of risks for his or her risk group, and compute the optimal coverage choice from the offered menus of short- and long-term disability coverage.

---

<sup>17</sup>Fortunately in this regard, the simulated correlations remain about the same when we increase risk aversion by a factor of 10 (maintaining the same coefficient of variation). The correlation drops somewhat when risk aversion is decreased by a factor of 10, thus making the qualitative results we report fairly robust to the calibrated level of relative risk aversion.

Using this model we simulated choices from the modal short- and long-term disability menus offered in the data, and correlated these choices with each other and with the underlying (simulated and therefore known) coefficient of risk aversion.



Table 1: Employee characteristics in baseline sample

	Mean	Std. Dev.	5th pctile	95th pctile
<b>Panel A: Demographics</b>				
Age	43.9	9.2	28	58
Annual wage (000\$)	58.4	71.7	25.6	114
Job tenure with Alcoa (years)	13.2	9.6	1	30
Female	0.23			
White	0.85			
Hourly (non-salary) employee	0.32			
Unionized employee	0.02			
Single coverage tier <sup>a</sup>	0.19			
Number of covered individuals per employee <sup>a</sup>	2.92	1.46	1	5
<b>Panel B: Annual Payouts by domain</b>				
Health insurance claims (\$)	5,221.4	10,606.8	60.3	18,091.7
Prescription drug insurance claims (\$)	1,491.8	2,162.2	0.0	5,507.3
Dental insurance claims (\$)	781.3	837.3	0.0	2,443.0
Short-term disability insurance (fraction with any claims) <sup>b</sup>	0.061			
Long-term disability insurance (fraction with any claims) <sup>c</sup>	0.002			
Annual 401(k) contribution (\$)	4,616.2	3,199.5	709.6	11,225.8

The table is based on the 12,752 employees who constitute our baseline sample.

<sup>a</sup> The coverage tier and covered individuals are based on the medical coverage choices; we view them as reasonable proxies for family size and structure.

<sup>b</sup> Conditional on having a short-term disability claim, the average claim length is 51 days.

<sup>c</sup> Conditional on having a long-term disability claim, the average claim length in our data is 345. However, the long-term claim data is truncated at about two years, so 345 should be viewed as a lower bound.

Table 2: Summary of benefit options

	Share	Premium saving relative to safest option	Expected incremental cost	Std. Dev. Of incremental cost
	(1)	(2)	(3)	(4)
<b>Health Insurance</b>				
Option 1	17.3%	1,016.6	1,415.6	1,052.4
Option 2	1.3%	747.7	880.0	559.7
Option 3	2.7%	545.3	645.6	380.8
Option 4	26.3%	325.0	350.8	173.4
Option 5	52.4%			
<b>Prescription Drug Insurance</b>				
Option 1	23.8%	181.2	248.6	385.0
Option 2	9.7%	109.6	124.3	192.5
Option 3	66.4%			
<b>Dental Insurance</b>				
Option 1	30.0%	95.7	45.2	112.9
Option 2	70.0%			
<b>Short-Term Disability Insurance<sup>a</sup></b>				
Option 1	15.5%	165.1	140.2	825.7
Option 2	17.9%	63.5	70.3	413.4
Option 3	66.6%			
<b>Long-Term Disability Insurance<sup>a</sup></b>				
Option 1	16.3%	152.4	17.0	395.7
Option 2	14.9%	63.5	8.5	197.9
Option 3	68.8%			
<b>401(k) allocation<sup>b</sup></b>				
Risk-free 0%	40.6%	--	-421.7	514.0
Risk-free 0-25%	19.9%	--		
Risk-free 25-50%	12.8%	--		
Risk-free 50-75%	6.5%	--	-210.8	257.0
Risk-free 75-100%	3.4%	--		
Risk-free 100%	16.8%	--		

All options are shown in the ordinal ranking from more (option 1) to less risk exposure (with the possible exception of health insurance option 1; see text and Appendix Tables A1 and A2 for details). Column (1) shows the fraction who chose each option in our baseline sample. Column (2) shows the average (in the baseline sample) premium savings from choosing a given option relative to choosing the safest (least risk exposure) option; these vary across employees based on benefit menu, coverage tier (for health, drug and dental), and wages (for short- and long-term disability). Columns (3) and (4) show, respectively, the average and standard deviation of the incremental cost that the insurer would face (counterfactually for most of the sample) in covering our baseline sample of employees, given the realized spending and coverage tier choices, with the safest option (i.e., the highest numbered option) relative to the option shown.

<sup>a</sup> Short-term and long-term disability benefits (columns (3), and (4)) and premiums (column (2)) are proportional to the employee's wage.

<sup>b</sup> For 401(k), columns (3) and (4) report expected incremental dollar payout (and associated standard deviation) for 0% vs. 100% in risk-free asset (first row) and 50% vs. 100% in risk-free asset (second row) assuming the average annual employee contribution in our baseline sample of \$4,616. For the risky investment portfolio, we assumed the allocation across different risky funds observed in the baseline sample, and similarly for the risk free part of the investment portfolio (see Table A2).

Table 3: Main results

**Panel A: Spearman rank correlations**

	Health	Drug	Dental	STD	LTD
Drug	0.400				
Dental	0.242	0.275			
STD	0.226	0.210	0.179		
LTD	0.180	0.199	0.173	0.593	
401(k)	0.057	0.061	0.036	0.029	0.028
				(0.002)	(0.002)

Average correlation is 0.192

**Panel B: Baseline specification, multivariate regression**

	Health	Drug	Dental	STD	LTD
Drug	0.452				
Dental	0.238	0.267			
STD	0.188	0.197	0.169		
LTD	0.155	0.191	0.165	0.600	
401(k)	0.057	0.056	0.035	0.029	0.018
				(0.001)	(0.042)

Average correlation is 0.188

**Panel C: Multivariate regression, Controlling for predicted and realized risk**

	Health	Drug	Dental	STD	LTD
Drug	0.412				
Dental	0.207	0.25			
STD	0.155	0.156	0.156		
LTD	0.129	0.156	0.153	0.593	
401(k)	0.039	0.032	0.026	0.002	-0.002
			(0.004)	(0.844)	(0.817)

Average correlation is 0.164

The table reports results for our baseline sample of 12,752 employees. Unless reported otherwise in parentheses, the p-values associated with whether the correlation coefficient is different from zero are all less than 0.001. Each cell reports a pairwise correlation. The average correlation is simply the average of the fifteen pairwise correlations shown, and is provided only as a single summary number. Panel A reports Spearman rank correlations, and Panel B reports the correlation structure from the multivariate regression shown in equation (1) with control (indicator) variables for the benefit menu the employee faces. Panel C reports the results from another variant of equation (1) which additionally includes controls for predicted and realized risk in all domains for each equation; see the text for details on the construction of these risk variables.

Table 4: Predictive power of different variables

Regressors	Dependent variable					
	Health	Drug	Dental	STD	LTD	401(k)
Choices in other domains	0.227	0.243	0.102	0.374	0.368	0.004
Predicted and realized risk	0.067	0.106	0.056	0.043	0.023	0.023
Demographics	0.037	0.044	0.025	0.039	0.033	0.043
Choices in less related domains	0.082	0.102	0.077	0.063	0.054	0.004
All of the above	0.245	0.292	0.144	0.394	0.378	0.046

Each entry in the table reports the adjusted  $R^2$  from a separate OLS regression of the dependent variable shown in the column heading. In all regressions, the dependent variable is the enumerated coverage choice in the domain given by the column header, after partialing out menu fixed effects. The regressors are given by the row header. “Choices in other domains” contain the vector of the enumerated choices in all five other domains. “Predicted and realized risk” refers to a vector of both predicted and realized risks in all domains (see text for more details on how these are constructed). “Choices in less related domains” omits the other choice which is most correlated with the dependent variable (Drug in Health and Health in Drug, Drug in Dental, LTD in STD and STD in LTD, Health in 401(k)). Demographics consist of age, age squared, dummy variables for gender, race and employee type (hourly or salary), job tenure in Alcoa, annual wage, and a dummy for single coverage tier (as a proxy for family composition).

Table 5: Summary correlations by groups

	Obs.	Average correlation	Health-Drug correlation	Health-STD correlation	Health-401(k) correlation
	(1)	(2)	(3)	(4)	(5)
(1) Single coverage	2441	0.224	0.532	0.252	0.074
Non single	10311	0.176	0.421	0.167	0.055
(2) More tenured	11708	0.185	0.448	0.184	0.059
Newly hired	1044	0.195	0.472	0.184	0.023
(3) Higher wage	3151	0.178	0.425	0.146	0.072
Lower wage	3173	0.162	0.439	0.174	0.026
(4) Don't allocate to Alcoa Stock	7468	0.193	0.448	0.195	0.073
Allocate to Alcoa stock	5284	0.180	0.456	0.176	0.033
(5) Rebalance 401(k) portfolio	3626	0.186	0.430	0.178	0.079
Don't rebalance	9126	0.188	0.460	0.190	0.049
(6) Over 55 years old	1700	0.167	0.446	0.147	0.061
Under 35 years old	2568	0.199	0.447	0.209	0.031
(7) Salaried employees	8644	0.187	0.442	0.175	0.069
Hourly employees	4108	0.157	0.453	0.170	0.016

The table reports the correlation coefficients for the subsamples specified in the row headers. The estimates all use our baseline specification (Panel B of Table 3). That is, we report the correlation structure of the error term from estimating the multivariate regression shown in equation (1) with covariates for benefit menu fixed effects. The average correlation in column (2) is the simple average across the fifteen possible pairs of correlations (as in the bottom of each panel of Table 3), while the other columns report the pairwise correlations for the selected pairs shown in the column headings. Row 1 divides the sample by single coverage tier for health and drug vs. all other (non-single) coverage tiers. Row 2 separates out newly hired employees (defined as less than 2 years of tenure) from higher tenured employees. Row 3 separately examines employees with greater than \$72,000 annual wages and less than \$36,000 annual wages (approximately the top and bottom quartiles of wages). Row 4 separates employees who did and did not allocate their own 401(k) contributions to Alcoa stock. Row 5 separates employees who did and did not rebalance their 401(k) portfolio during the year.

Table 6: Robustness

	Obs.	Average correlation	Health-Drug correlation	Health-STD correlation	Health-401(k) correlation
	(1)	(2)	(3)	(4)	(5)
1 Baseline specification	12,752	0.188	0.452	0.188	0.057
<b><u>A. Alternative specifications</u></b>					
2 A system of ordered probits	12,666	0.264	0.550	0.292	0.055
3 Discretizing the 401(k) choice	12,752	0.189	0.452	0.188	0.058
4 Control for coverage tier	12,752	0.186	0.447	0.187	0.058
5 Use only the largest pricing menu	7,722	0.195	0.452	0.191	0.069
<b><u>B. Alternative samples</u></b>					
6 Include those in opt-out and HMO	15,409	0.165 <sup>a</sup>	--	--	--
7 Include employees who did not contribute to 401(k)	15,402	0.257 <sup>b</sup>	0.446	0.184	--
8 Include those not offered LTD coverage	15,675	0.162 <sup>c</sup>	0.442	0.183	0.052
9 Exclude those in Health Option 1 (due to HRA component)	10,547	0.147	0.226	0.175	0.009
10 Include only new hires	1,044	0.195	0.472	0.184	0.023
11 Exclude individuals who may have chosen default options	11,323	0.191	0.460	0.197	0.059

This table reports correlation results for variants of the baseline specification. Column (2) shows the simple average of the 15 pairwise correlations, and columns (3) through (5) report correlations for specific pairs. Row 1 replicates the baseline specification (as in Table 3, Panel B) which reports the correlation coefficients of the error term from the multivariate regression shown in equation (1), including benefit menu fixed effects. All rows except row 5 include these benefit menu fixed effects. Row 2 estimates a system of five ordered probits and one linear regression (for the 401(k) domain) rather than the multivariate regression shown in equation (1) (see text for more details). Row 3 replaces the continuous 401(k) riskiness measure with a discretized ordinal measure of 1-5, for each 20% interval of allocations in safe assets. Row 4 includes coverage tier (based on health coverage) fixed effects, and row 5 reports results using the largest (modal) benefit menu (and therefore no menu fixed effects). Rows 6-11 report results from alternative samples. In rows 6, 7, and 8 we include employees that were excluded from the baseline sample, and in these cases we omit the domain that had disqualified these employees from the baseline sample. Therefore the average correlations in these cases are not directly comparable to the baseline specification, although the individual pairs are. In row 10 we limit the sample to new hires (defined as job tenure at Alcoa of less than two years). In row 11 we exclude the approximately 10% of the employees whose choices are fully consistent with the default options in all insurance domains, and are therefore potentially “passive” choosers.

<sup>a</sup> The comparable average correlation (that is, over the 6 pairs that do not include health and drug coverage) in the baseline specification is 0.169.

<sup>b</sup> The analogous average correlation (that is, over the 10 pairs that do not include 401(k) choices) in the baseline specification is 0.262.

<sup>c</sup> The analogous average correlation (that is, over the 10 pairs that do not include long-term disability coverage) in the baseline specification is 0.169.

Table A1: Coverage Details for Insurance Plans

Summary of Key Coverage Details (1)		Additional details (2)
Health Insurance <sup>a</sup>	Deductible (In-network / out-of-network)	
Option 1 <sup>b</sup>	3,000 / 6,000	
Option 2	1,500 / 3,000	After satisfying the annual deductible, cost sharing is 10% in-network and 30% out-of-network for all options. All options also specify in-network and out-of-network out-of-pocket maximums, but these are rarely binding. Preventive care is covered in full under all coverage options.
Option 3	1,000 / 2,000	
Option 4	500 / 1,000	
Option 5	0 / 500	
Prescription Drug Insurance	Cost sharing for branded drugs (retail / mail order)	
Option 1	50% / 40%	All options have cost-sharing of 10% for generic (non-branded) mail order drugs and 20% for generic retail drugs. All options have a \$50 deductible (\$100 for family) and a \$50 (\$100 for mail-order) maximum per prescription.
Option 2	40% / 30%	
Option 3	30% / 20%	
Dental Insurance	Per person Deductible / Maximum annual benefit	The family deductible is double the per-person amount. Both plans fully cover preventative care, provide identical coverage for other special treatments. Oral surgery is covered at 50% under option 1 and 100% under option 2. Orthodontia is not covered under option 1 and is covered at 50% under option 2.
Option 1	50 / 1000	
Option 2	25 / 2000	
Short-Term Disability Insurance <sup>c</sup>	Wage replacement rate	Salary workers have 100% replacement rate for first two weeks of disability under all options; all options provide up to 26 weeks of benefits.
Option 1	mostly 60% (sometimes 40%)	
Option 2	mostly 80% (sometimes 60%)	
Option 3	mostly 100% (sometimes 80%)	
Long-Term Disability Insurance <sup>c</sup>	Replacement rate	All long-term disability coverage is payable after 26 weeks of disability (when the short-term disability coverage is capped).
Option 1	mostly 50%	
Option 2	mostly 60%	
Option 3	mostly 70%	

All options are shown in the ordinal ranking from more (option 1) to less risk exposure (with the possible exception of health insurance option 1; see note b and text for details). Column 1 summarizes key features of each option. Column 2 provides additional details.

<sup>a</sup> Health insurance: deductibles are shown for the non-single coverage tier; deductibles for single coverage are half what is shown.

<sup>b</sup> Option 1 includes a Health Reimbursement Account (HRA) in which Alcoa contributes \$1,250 in tax free money each year that the employee can use to fund eligible out of pocket health care expenses. Any balance remaining at the end of the year can be rolled over to pay for future out of pocket costs. See text for more details.

<sup>c</sup> Short-term and Long-term disability benefits (column (1)) are proportional to the employee's wage.

Table A2: List of funds available for 401(k) allocation

Fund name (Asset Class)	Monthly return				
	Share <sup>a</sup>	Mean	Std. Dev.	Min.	Max.
<u>Classified (by us) as "Risk Free":</u>					
GIC/Stable Value (Fixed Income)	24.47%	0.35	0.02	0.31	0.37
Vanguard Total Bond	3.95%	0.42	0.83	-1.09	1.92
<u>All other classified as risky:</u>					
American Balanced (Balanced Equity)	10.58%	0.65	1.36	-2.34	2.89
Inv. Co. of America (Large Cap US Equity)	9.62%	0.83	1.84	-3.82	3.86
AMCAP (Large US Equity)	6.77%	0.66	2.06	-4.19	4.01
Vanguard Institutional Index (Large Cap US Equity)	9.42%	0.79	2.21	-4.18	4.43
MSDW International Equity	4.09%	1.25	2.32	-3.30	4.92
New Perspective (International Equity)	5.34%	1.49	2.72	-4.13	6.32
Putnam OTC (Mid Cap US Equity)	3.23%	1.01	3.40	-6.35	7.45
Small Cap Core (Small Cap US Equity)	0.30%	0.29	3.44	-6.95	7.90
Putnam Vista (Mid Cap US Equity)	3.71%	0.56	3.55	-8.58	6.75
MSDW Emerging Markets	2.62%	3.13	5.83	-11.69	15.03
Company (Alcoa) Stock Fund	15.90%	1.30	6.71	-8.85	16.79
<u>Benchmarks during the same period:</u>					
Risk free <sup>b</sup>	--	0.37	0.05	0.26	0.43
S&P 500	--	0.63	2.21	-4.40	4.33

Employee contributions to their 401(k) accounts can be made with either pre- or after-tax dollars. Employees can contribute 1-16% of eligible pay with some additional restrictions for some highly paid employees. In our sample, Alcoa usually matches 100% of pre-tax contributions, up to 6% of eligible pay. Employer (Alcoa) contributions are always invested in the company stock and can only be moved to a different fund after two years. In the 2004 data that we are using, the above 13 funds are available for contributions (sorted by the standard deviations of monthly returns). In the analysis we use as a measure of riskiness of the portfolio the share of employee contributions invested in those (two) funds that are presented as least risky. Indeed, as apparent from the table, these two funds exhibit less volatility (and mostly lower expected return). Employees also have the option to invest in a personal choice retirement account in which they have access to other funds besides the 13 funds just described. Direct contributions to this fund are not possible, only transfers, and we do not have detailed data on the composition of investments in these funds. For our analysis we only use direct employee contributions. In 2004 only about 28 percent of the sample rebalances and 24 percent of the sample changes the allocation of their contributions. The average employee contribution in the baseline sample (which restricts attention to non-zero contributions) is around \$4,600. About 40 percent of the sample has no contributions to the risk free funds, and about 17 percent invest all their contributions in the risk free funds. Just over 40 percent of the sample has some employee contributions invested in company stock. The series of returns are based on monthly returns over the 29 month period from August 2005 to December 2007, which was the longest time period for which we have consistent returns data for all funds. Returns data are from CRSP (when available), or from Hewitt (when CRSP data are not available, for the few funds that are not publicly traded).

<sup>a</sup> We compute the share of dollars contributed to each fund out of total 401(k) contributions made by all employees in our baseline sample.

<sup>b</sup> For the risk free benchmark we use the CRSP three month Fama Risk Free Rates series, which are derived from average lending and borrowing rates.