# How general are risk preferences? Choices under uncertainty in different domains

# How general are risk preferences? Choices under uncertainty in different domains*

Liran Einav, Amy Finkelstein, Iuliana Pascu, and Mark R. Cullen†

**Abstract.** We analyze the extent to which individuals' choices over five employer-provided insurance coverage decisions and one 401(k) investment decision exhibit systematic patterns, as would be implied by a general utility component of risk preferences. In particular, we examine whether individuals display a stable ranking in their willingness to bear risk, relative to their peers, across the different contexts. We provide evidence consistent with an individual-specific but domain-general component that operates across all of the choices. This component appears substantial among the five insurance domains; 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 larger for individuals who may be associated with better "financial sophistication." Estimates from a stylized coverage choice model suggest that up to 30 percent of our sample makes choices that may be consistent across all six domains.

*JEL classification numbers*: D14, D81, G11, G22

**Keywords:** Risk aversion, Insurance, Uncertainty, Portfolio choice

---

*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, Ben Handel, Glenn Harrison, David Laibson, Dan Silverman, Jon Skinner, three anonymous referees, and Penny Goldberg (the editor) 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), the John D. and Catherine T. MacArthur Foundation Network on Socioeconomic Status and Health, and Alcoa, Inc. (Cullen), and the U.S. Social Security Administration. This research was supported by the U.S. Social Security Administration through grant #10-M-98363-1-02 to the National Bureau of Economic Research as part of the SSA Retirement Research Consortium. The findings and conclusions expressed are solely those of the author(s) and do not represent the views of SSA, any agency of the Federal Government, or the NBER.

†Einav: Department of Economics, Stanford University, and NBER, leinav@stanford.edu; Finkelstein: Department of Economics, MIT, and NBER, afink@mit.edu; Pascu: Department of Economics, MIT, iuli@mit.edu; Cullen: Department of Internal Medicine, School of Medicine, Stanford University, 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 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 unique data on thousands of individuals and analyzing actual decisions that each of them make regarding financial lotteries in different domains.

Specifically, we examine the workplace-based benefit choices that Alcoa employees make concerning their 401(k) asset allocations, their short-term disability insurance, their long-term 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 or idiosyncratic preferences for a given physician. 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 expected risks of similar magnitudes, making decisions across contexts more comparable.

Our focus is on quantifying the empirical importance of any individual-specific, domain-general component of preferences rather than on testing the extreme nulls of complete consistency or no consistency in preferences across domains. Neither extreme null strikes us as particularly compelling in practice; reality almost surely lies in between. Perhaps more importantly, as we discuss in more detail below, while it seems possible to plausibly 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 additional difficult-to-test modeling assumptions.

A key challenge that we face in developing an approach to quantifying the extent of domain generality of preferences, is that in our interest to examine the stability of preferences across contexts, we would like to avoid *context-specific* modeling assumptions that could push us toward one finding or another. A natural way to evaluate the stability of risk preferences across domains would be to write down a model of consumer behavior, use the data and the model to obtain estimates for risk aversion for each individual in each domain, and then compare these estimates. Cohen and Einav (2007) provide a framework for inferring risk aversion from insurance choices, which could be adapted to our various contexts. However, their framework also illustrates that estimating the distribution of risk aversion from individuals' insurance choices involves a number of domain-specific modeling assumptions regarding the nature of ex-ante information, expectation formation, the risk realization process, the nature of heterogeneity in risk and risk preferences, the possibility of moral hazard, and the class of utility functions. While these assumptions are not a problem per se, in assessing the extent of domain generalizability of preferences, one would naturally worry greatly about the role of domain-specific modeling assumptions. Given this challenge, in this paper we shy away from specifying a complete model of primitives for each domain. Instead, we pursue two other complementary strategies that allow us to make progress in investigating the motivating question while trying to minimize the need for domain-specific modeling assumptions.

Our first strategy takes a "model free," statistical perspective. We avoid any economic modeling of primitives and instead focus 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. Our results reject the null hypothesis that there is no domain-general component of preferences: individuals' choices across domains are positively correlated. More interestingly, in our view, we develop several benchmarks 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 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 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 the predictive power of one's 401(k) portfolio choice for one's insurance choices is systematically greater for individuals who are older, have more experience within the firm, have higher income, or who appear to be more financially sophisticated (as measured by external proxies in the data). This suggests that such individuals may fit better the canonical model.

The advantage of this "model free" approach is that it allows us to make inferences that are much more robust to various assumptions. In particular, the approach only requires us to assume that any unobserved individual- and domain-specific components 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 or on the way in which individuals form expectations, weight probabilities, and so on. The drawback of a "model free" approach is that the results cannot be directly mapped to underlying economic primitives. While we attempt to develop several benchmarks that may help in assessing whether the correlations we find point to a greater or lesser importance of the domain-general component of preferences, one can reasonably argue that such benchmarks are somewhat ad hoc.

Indeed, our second empirical approach attempts to link our results to underlying economic primitives. In particular, we estimate the fraction of our sample that makes choices across domains that can potentially be rationalized with a common risk aversion parameter. We write down a stylized model, which allows us to use the same (stylized) model across the different domains. This strategy trades off the need for a model-based framework with the concern mentioned above regarding too many domain-specific modeling assumptions. The key decision in this respect is – as in our first strategy – to focus on comparing the ranking of risk aversion rather than the levels. We do so in our second approach by allowing for a domain-specific (but constant across individuals) parameter, which essentially frees up the level of risk aversion in any context. While this minimizes the number of domain specific assumptions, it still requires us to make some assumptions that were not needed for our first, "model free" approach.

Our baseline results suggest that, subject to a domain-specific (but not individual specific) free parameter, just over 30 percent of our sample make decisions that could be rationalized across all six domains. This result appears robust to a number of variations to our baseline specification. In addition, we once again find evidence suggesting that preferences are less consistent across "less close" domains, particularly between the 401(k) asset allocation and the other five, insurance domains.

Overall, we view our findings from the two complementary approaches as generally supportive of a fair amount of domain generality in decision making under uncertainty. We should recall, however, our decision to focus on the stability across contexts in the relative ranking of individuals' risk preferences, rather than the stability of the absolute level of risk aversion. While appealing in reducing the necessary assumptions we need to make, this decision also makes it a more modest test of the canonical model. For example, our findings of a reasonable degree of consistency in individuals' relative ranking of risk preferences across domains does not preclude a rank preserving difference in the entire distribution of willingness to bear risk across domains. In addition, our findings of higher correlation in risk preferences across "closer" contexts suggests that our findings of quantitatively meaningful domain generality may not persist if we looked at more disparate contexts than those studied in this paper. We return to this briefly in the conclusion.

Our study is not alone 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.[1] Several studies have addressed the stability of risk preferences by investigating individual responses to financial lotteries across

[1] 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.

different types of lotteries and over time (Choi et al., 2007; Andersen et al., 2008a; Kimball, Sahm, 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.[2] A recent study by Dohmen et al. (2011) is probably the closest of this literature to our first approach; somewhat similar to Barsky et al. (1997), Dohmen et al. (2011) use a large data set of survey responses to hypothetical financial lottery questions and validate 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, our study is based on actual market choices. By contrast, many of the existing studies rely on individual responses to hypothetical questions (e.g., Barsky et al., 1997) or to self-reported behaviors (e.g., Cutler and Glaeser, 2005). 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. An approach that circumvents many of these concerns is the use of lab experiments with real consequences associated with the choices (e.g., Choi et al., 2007) or field experiments with a representative sample of a population, again involving choices with real (and non trivial) payoffs (e.g., Andersen et al., 2008a). Nonetheless, as Harrison et al. (2007) nicely show, mapping choices made in the lab to choices made in naturally occurring settings is not at all straightforward. This distinction makes it important to combine data from inside and outside the lab, either within the same paper as in Andersen et al. (2008b) or across papers, to which the current study contributes.

We are aware of only one other study of the stability of risk preferences across contexts that uses actual market outcomes. Barseghyan, Prince, and Teitelbaum (forthcoming) and Barseghyan et al. (2010) 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, to test whether they can reject the null that risk aversion is completely general across domains, and to explain the deviations they find using a non-expected utility framework. Our second approach is quite similar to theirs. It differs primarily in its scope – we look at a much broader, and less similar, range of domains – and, relatedly, in its focus and empirical approach. Barseghyan, Prince, and Teitelbaum (forthcoming) 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. By contrast, we focus on quantifying (rather than testing) the extent of domain generality in risk preferences after allowing for a domain-specific free component of risk preferences. Their approach is

[2] See also Chabris et al. (2008) for a similar exercise that focuses on discount rate (rather than on risk preferences).

a more ambitious one but relies on commensurately greater context-specific modeling assumptions, which are less troublesome in their more closely related domains. We therefore view the papers (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. (2011) – is our attempt to quantify the magnitude of any domain-general component of preferences by benchmarking it against plausible 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 "model free" approach and correlation results concerning the stability in individuals' relative ranking of risk preferences across contexts. Section 4 presents estimates from our second, model-based approach regarding the fraction of individuals whose choices may be rationalized across domains. Section 5 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 decision regarding the 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 a pre-specified choice of investment allocations, which in principle could be adjusted at any given time (although in practice only about one quarter of the 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 risks 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.[3]

**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 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). This 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.[4] 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.[5] 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 options 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

[3] 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%).

[4] 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).

[5] 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.

more detail.

**Description of coverage options.** As mentioned, we investigate employees' choices over six different domains. Table 2 summarizes the key features of each domain, with the options enumerated within each domain (as presented 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.[6] These options only vary in their financial coverage, and (with the exception of option 1) are vertically rankable,[7] with the deductible level being the key difference.[8] 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

[6] 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.

[7] 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 at 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. In the robustness analysis below we verify that results are 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.

[8] 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.

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 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.[9] 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.[10] The most interesting point we take away from Table 2 is that the incremental decisions across each domain are quite comparable in their expected magnitude, with incremental (annual) premiums (and associated benefits) ranging from several hundred to a few

[9] 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.

[10] 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. As long as an individual believes there is a sufficiently low probability of the relevant claim, he may not prefer the safest option.

thousand dollars. Of course, the overall magnitudes of the underlying risks can vary vastly (e.g., between long-term disability and dental), but the incremental coverage – which is the key for the coverage choice – is of a much more similar magnitude across domains.

**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 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 the canonical model of decision under uncertainty.

Third, as discussed earlier (and shown in Table 2, columns (2) and (3)), the choices within a domain are over similarly sized risks.

Fourth, many of the benefit options are entirely new in 2004, and the old options were no longer available. This means that for these benefit options we are looking at decisions made 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.[11] Specifically, the health, drug and dental options were all completely new – the old options were no longer available in these domains – while the disability options remained the same but their prices changed; the 401(k) options did not change.[12] As a further check against the possibility of "stale" decisions (particularly for 401(k)

[11] Given the substantial evidence on inertia in insurance choices (see Handel (2010) for a recent example) we would worry greatly about examining choices that may have been made a long time earlier (when an individual's characteristics may be different from what we currently observe) and/or at different times for different products.

[12] 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

allocations and potentially disability choices), we show in our robustness analysis that 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 some 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." Of course, the range spanned by our choices is much narrower than the full set of decisions under uncertainty that individuals make; in the end of the paper we discuss some of the challenges in extending the study to a broader range of domains.

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 relatively rich demographic information.

# 3 A "Model Free" Approach

## 3.1 Empirical Strategy

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 in individual's choices from among the (vertically ranked) options in each domain. We thus 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 from a system of six

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.

equations of the form:

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

where $x_i$ is a vector of control variables (which is the same in all equations in the system of equations), $\beta$ is a vector of domain-specific coefficients, and the main object of interest is the correlation matrix of the residuals.

We estimate this system in two separate ways. We first treat each equation as an ordered probit specification (except the 401(k) equation, which is treated as a regular equation with a continuous dependent variable) – that is, we assume that the six residuals are drawn from a multivariate normal distribution, and that the dependent variable is a latent domain-specific variable that maps a one-dimensional index into a discrete ordered coverage choice.[13] This specification treats properly the ordinal nature of the choices, but has the disadvantage that it does not lend itself to a natural $R^2$ measure, which we use later to compare the predictive power of different variables. We therefore also estimate the system of equations above using multivariate least squares, by enumerating the choices from 1 to $n$ in each domain (as in Table 2), and assigning them a cardinal interpretation despite their ordinal nature. This specification does not require us to assume that the errors are distributed normally and, more importantly, makes it natural to use $R^2$ to compare results across different specifications. As we report below, the correlation results that we obtain from the three specifications – the rank correlation, the system of ordered probits, and the multivariate regression analysis – are all very similar.

Because standard theory models insurance choices as driven by risk and risk aversion, it is essential to control for risk if one wants to make inferences about risk aversion. The baseline set of control variables ($x_i$) we include in the ordered probit and multivariate least squares specifications are dummy variables for the menu of benefits the employee faced (described above). We also explore the sensitivity of our results to the inclusion of additional controls (in all six equations) that proxy for individual risk in each of the five insurance domains. We attempt to control for two components of risk; the first is risk that can be predicted using observables, and the second is an individual-specific risk component, which is idiosyncratic to the individual.

To proxy for the predictable component of risk, we use two measures. The first measure is based on a statistical model of realized risk in each domain on a flexible functional form of our observables; we generate and then use as controls the model predictions.[14] A second measure of

[13] We estimate this 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 and Roodman (2009).

[14] The results are not at all sensitive to the precise way we predict risk. For the results we report below, risk is predicted from a linear regression of realized risk (dollar spending for health, drug, and dental insurance; and days of

predictable (health) risk is based on software that predicts future medical spending on the basis of previous years' detailed medical diagnoses and claims, as well as demographics.[15] To proxy for the idiosyncratic component of risk we use the realization of that risk in the subsequent coverage period. That is, if individual risk is realized from an individual-specific distribution, conditional on observable risk, the realization of risk can be used as a (noisy) proxy for the underlying ex-ante individual-specific risk type. The identification arguments in Cohen and Einav (2007) and in Einav, Finkelstein, and Schrimpf (2010) use a similar idea. Finally, to allow for correlation in both observed and unobserved heterogeneity in risk across domains, we include controls for *all* our proxies in *all* the insurance domains. That is, each equation includes eleven control variables, containing predicted and realized risk in each of the five insurance domains, as well as the software-generated prediction of health risk.

### 3.2 Results

Table 3(a) presents the baseline correlation results, when we do not use additional control variables (except for benefit menu fixed effects in Panels B and C). 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 correlations, as a single summary measure. Panel B shows the estimated correlation from the system of ordered probit specifications (and a single 401(k) linear equation), and Panel C shows the correlations from the baseline multivariate regression described above. In general, we can (easily) reject the null hypothesis of a correlation of zero.

By rejecting the null hypothesis of a correlation of zero, 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. Although the finding that risk preferences are correlated across domains may be viewed as hardly surprising, from the perspective of the canonical model, it is encouraging to find this positive correlation so robustly across a broad range of contexts.

This test of the admittedly not very compelling null of no domain-general component of choices is subject to the important caveat that non-preference factors may introduce correlations across domains. In the case of insurance, a natural suspect is potential 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

disability for either disability insurance) on: (i) cubic splines for age, wage, and job tenure; (ii) dummy variables for gender, race, employee type (hourly or salary), union status, single coverage for health benefits, family size, and state fixed effects; and (iii) interaction variables between age and the gender, employee type, and single coverage dummy variables.

[15] This is a relatively sophisticated way of predicting medical spending as it takes into account the differential persistence of different types of medical claims (e.g., diabetes vs. car accident) in addition to overall utilization, demographics, and a rich set of interactions among these measures. The particular software we use is a risk adjustment tool called DXCG risk solution which was developed by Verisk Health (http://www.veriskhealth.com/) and is used, e.g., by the Center for Medicare and Medicaid services in determining reimbursement rates. See Carlin and Town (2010) and Handel (2010) for other examples of academic uses of this type of predictive diagnostic software.

most compelling context to test the null of complete domain specificity.

To try to address the concern about underlying risk correlations across insurance domains, Table 3(b) reports the analogous results after we add control variables (as explained earlier) for both predicted and realized risk in *all* domains in *each* equation. Panel A reports results from the specification of a system of ordered probits and Panel B for the multivariate regression.[16] The results, again, are very similar across the two specifications, and quite remarkably the magnitude of the correlations generally remains almost the same as in Table 3(a), with only a slight decline (the decline is to be expected, given that the risks are positively correlated across domains). 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 fully specified economic models – 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, forthcoming).

Across all panels of Table 3(a) and Table 3(b) we see that the average pairwise correlation is 0.16 to 0.26. Perhaps not surprisingly, there is a pronounced pattern of substantially higher correlation coefficients between pairs that are more "similar." For example, in panel B of Table 3(a), the correlation between drug and health coverage choices is 0.55 and the correlation between long-term and short-term disability insurance choices is 0.77. By contrast, health insurance and short-term disability insurance show only a 0.29 correlation and the lowest pairwise 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 0.07 or less). 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.

We also examine how the correlation in choices varies across different identifiable groups. Table 4(a) and Table 4(b) present the main results for the ordered probit and multivariate specifications, respectively. Specifically, the results show 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 4 is in column (5), which shows 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. 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).

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 aver-

[16] Table 3(b) does not report the Spearman rank correlations, for which it is less obvious how to add controls.

sion 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 risk perceptions or in the mapping from "true" underlying risk preferences to choices, for individuals who appear to be more "financially literate"; such an interpretation could suggest that the correlation results underestimate the importance of the domain-general component of risk preferences in the full sample. 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 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.

### 3.3 Robustness

We explored the robustness of our main correlation results (Table 3(a), Panels B and C) to various alternative specifications and samples. Tables 5 and 6 summarize the results of these analyses. As in Table 4, in the interest of space, we do not report every pairwise correlation, but instead report the average correlation and the correlations of three selected pairs. We explore two main types of sensitivity analysis: alternative specifications and alternative samples. Unless otherwise specified, each row represents a single change relative to the baseline specification. Overall, the results seem to be quite robust to the alternative exercises we explore.

**Alternative specifications and sample definitions.** Table 3(a) already showed that the Spearman rank correlations, the correlations estimates that are based on the system of ordered probits, and the linear multivariate regression all lead to similar results. Table 3(b) has also shown that the results are not affected much by the inclusion of a large set of controls for risk. Row 1 of the two panels of Table 5 replicates the baseline results (Table 3(a), panels B and C, respectively), and the rest of the rows in the table examine additional plausible concerns.

Row 2 examines a concern that perhaps the reason that the 401(k) choice is less correlated with all other insurance choices is driven by the fact that all insurance choices are discrete and ordinal, while the 401(k) choice is continuous and has a cardinal interpretation. To investigate this further, we discretize the 401(k) asset allocation decision and turn it into an ordinal measure, so it is more similar in nature to the other choices. We do so by taking the (continuous) measure of the percentage of employee contributions allocated to the safe funds, and convert it to a discrete integer between 1 to 3, with 1 corresponding to investing nothing in the safe funds, 2 corresponding to investing something but not everything in the safe funds, and 3 corresponding to investing everything in the safe funds.

In row 3 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 family coverage), and in row 4 we investigate concerns about whether our benefit menu fixed effects fully capture differences in choices due to prices by limiting the sample to those who faced the prices in the single largest benefit menu (about 60 percent of our baseline sample).

The rest of the rows in Table 5 explore the sensitivity of our baseline specification to alternative sample definitions. In rows 5 through 7 we add back in various employees who were excluded from the baseline sample. In row 5 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 6 we include employees who did not contribute to their 401(k) plan in 2004, and in row 7 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 5, 401(k) in row 6, and long-term disability in row 7). 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 8 we exclude from our analysis individuals who chose health insurance option 1, the lowest coverage option. As mentioned in Section 2, this option is bundled with a Health Retirement Account component, so it is not fully vertically rankable. In row 9 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 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 10 we exclude the roughly 11 percent of the individuals who might have been "passive" choosers, given that all their coverage decisions in the insurance domains were consistent with the default options.

**Outside insurance and investment choices.** A fundamental feature of our analysis is that while we have good data on individuals' decisions and outcomes within Alcoa, naturally we have very little information about any other of the individuals' insurance and investment portfolios, which are external to Alcoa. Thus, we may be missing important pieces of the overall insurance coverage for a particular risk, or the overall wealth portfolio. On the insurance front we are relatively sanguine. Given the generosity of Alcoa benefits relative to anything a spousal employer might provide, as well as the well-known problems with private markets (that are not employer-provided) for these insurance products, we think it is a reasonable approximation to assume that there is little non-Alcoa insurance purchase. However, non-Alcoa investments are a potentially important concern. To try to shed light on how important this may be for our results, we undertake two types of exercises.

First, to try to proxy for outside investments, we construct measures of the individual employee's housing wealth and then repeat our analysis by stratifying on housing wealth, so that we are comparing choices among individuals with relatively similar outside housing wealth. Of course, this strategy does not address other financial and non-financial wealth in the employee's portfolio. In practice, however, the retirement component is large relative to other financial assets for individuals with retirement financial wealth, and housing wealth is a very large share of non-financial wealth for such individuals (Bucks, Kennickell, and Moore, 2006). Therefore, controlling for housing wealth is likely a first order improvement in trying to address the non-Alcoa portfolio composition. To obtain data on housing, we matched the home addresses of our Alcoa employees to public records containing information on their home value and their equity stake in their house; we were able to link about one-third of our sample.[17]

The results are shown in Table 6. Once again we report estimates for the average correlation, the health-drug correlation, the health-short term disability correlation and the health-401(k) correlation. However, because this exercise may be particularly relevant for the sensitivity of the relationship between 401(k) choices and insurance choices, we also report each of the 401(k)-insurance product correlations. The first row shows results for the full sample, while the second row shows results for the sample for whom we were able to link in housing data ("housing subsample"). Rows 3-5 show results stratified (in roughly equally sized bins) by housing equity: less than \$50,000, \$50,000 to \$150,000, and above \$150,000. The results are not overly sensitive to this stratification. In particular, the basic pattern of much larger correlations among insurance choices than correlation between 401(k) portfolio allocation and insurance choices remains. The correlations are also extremely similar across employees with different equity levels or equity shares. For example, the health insurance-401(k) correlation is always lower than 0.07 for all across equity levels, while the correlation between health and drug coverage choices is always above 0.4. There does not seem to be any consistent pattern of a monotone relationship between housing equity and the magnitude of the various correlation coefficients.

Second, we tried to define a sample of employees who are less likely to have substantial non-401(k) financial investments by restricting the sample to employees who do not max out their possible 401(k) contributions; because of the favorable tax treatment of 401(k) investments, it seems plausible that individuals who are not saving as much as possible in tax preferred vehicles may have less outside savings than those who are. We therefore divide the sample into the approximately 14 percent who have contributed the maximum allowable amount to their 401(k) and the remainder who have not maxed out their allowable 401(k) contributions. The bottom two rows of Table 6 show that the results are broadly similar for the two groups. For example, the correlation between 401(k) portfolio allocation and insurance choice is slightly higher for those who have maxed out their 401(k) contributions for health insurance but slightly lower for the other four types of insurance. The general pattern of much larger correlations among insurance choices than between 401(k) portfolio allocation and insurance choices remains for both groups.

While of course these tests are limited in their nature, it is nonetheless reassuring to find that the results suggest that our inability to control for the entire wealth portfolio is unlikely to be

[17] The data were provided by a real estate data vendor DataQuick, which compiles data on real estate from public records such as county recordings of ownerships and transactions, and county tax assessors. See http://www.dataquick.com/sharedata.asp for more information. The employees for whom we were able to match housing data are unlikely to be a random sample of our employees; for example, we were unable to match employees with P.O. Boxes as addresses, and we likely have less success for counties without electronic records.

having a large impact on the correlations we examine.

### 3.4 Benchmarks

As noted at the outset, our primary interest is in developing reasonable benchmarks against which one can try to assess whether the correlation in the ordinal ranking of the riskiness of one's choices across domains suggests a quantitatively large or small domain-general component of risk preferences. Comparing the estimated correlations to the benchmark correlation of one does not provide a meaningful assessment of the extent of domain generality of preferences, or a test of the null of complete domain generality of preferences. We would not expect a rank correlation of one even if preferences were fully domain general.

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 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 an 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 (Gollier and Pratt, 1996; Guiso, Jappelli, and Terlizzese, 1996). This "background risk" problem introduces yet another reason why fully domain general preferences need not produce a rank correlation of one across domains.

Our first exercise that may allow us to start assessing whether the correlation estimates we report are large or small is to compare the predictive power of choices in other domains to the predictive power of demographic covariates. Table 7 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, which may make age a particularly important factor in 401(k) decisions.

A second exercise 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 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. In the multivariate regression, 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 multivariate correlation estimates across insurance domains reported in Table 3a, panel C, which range from 0.16 to 0.60.[18]

Our general conclusion from these benchmarks is that, contrary to our prior expectations, the reported average correlations of 0.16-0.26 are in fact quite high, and suggestive of an important domain-general component of risk preferences. To more specifically quantify the extent of domain generality of preferences requires that we link our results to underlying economic primitives. This in turn requires to move from a statistical model to an economic model, which is the focus of the next section.

# 4 A (Stylized) Model-Based Approach

## 4.1 A Calibration Exercise

We considered two (related) approaches to try to relate the statistical correlation in individual's choices across domains to underlying economic primitives, namely coefficients of risk aversion. One approach would be to start with a fully specified model of coverage choice, assume a benefit menu similar to the one observed in the data, and assume full domain generality (by imposing a common risk aversion parameter within an individual across domains). We could then simulate what the correlation coefficient between the implied coverage choices in different domains would be under this assumption of full domain generality, and compare it to what we have observed. This would allow us to obtain some benchmarks for the correlation coefficients between choices generated by a

[18] One could also investigate correlation in choices over time without any change in benefit design. The concern about such an exercise is that inertic behavior would be driving much of the results, 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.

model with fully domain-general risk preferences, but subject to the non-linearities and discreteness that arise because of the structure of the insurance options and the decision process.

In appendix A we describe such a calibration exercise, and apply it for short- and long-term disability choices. The model makes specific assumptions about the form of the utility function, about the expectations individuals have regarding their risks, and about the calibrated values of additional parameters such as the discount rate, the (common) distribution of risk aversion across individuals, and the nature of the risk realization processes. The results give rise to implied correlation coefficients between individuals' coverage choices of short- and long-term disability insurances. When we calibrate the average coefficient of relative risk aversion in the sample to 3, we find that the correlation between short- and long-term disability insurance range from 0.18 to 0.55, depending on the extent of correlation in risks we allows across domains. The range is somewhat larger (0.07 to 0.53) when we calibrate the average coefficient of relative risk aversion to 0.7. These reported ranges are just below the empirical correlation of about 0.6 for these two domains, as reported in Table 3.

Such an exercise is suggestive that the correlation coefficient we report in the data points to an important domain-general component of risk preferences. In principle, the robustness of the result could be probed with respect to the many underlying (strong) assumptions behind it. A more important concern is that the exercise would still be using only two specific domains. Although in principle, such an exercise could also be extended to additional domains, it is no coincidence that we chose two of the most similar domains for this exercise, so that the natural models for coverage choice were also quite similar

## 4.2 A Model

To extend the exercise to other domains and investigate robustness, we choose instead to pursue a second approach, which is in some sense the mirror-image of what we have just described. Instead of starting with a fully domain general model and asking what it would imply for the data, we start instead with the data and ask, in the context of a given, stylized model of coverage choice, what fraction of our sample's choices can be rationalized with a single (individual-specific) risk aversion coefficient. Our modeling approach is guided by a desire to reduce – although we cannot of course eliminate – domain-specific modeling assumptions. We therefore write down a stylized model of coverage choice that is stripped of many domain-specific details. This framework allows us to estimate the same generic model of primitives across the different contexts, which are quite different from each other. As we shall see, a key decision in this respect is to follow the spirit of our first "model free" approach by focusing on the (narrower) question of comparing the consistency of individual's ranking of risk aversion relative to their peers across contexts, rather than the consistency of individual's level of risk aversion across contexts.

Consider a domain $d$ and an individual $i$. We assume that choices are generated by expected utility maximizers who have a domain-invariant vNM utility function over wealth, $u_i(w)$. Faced with a set of coverage options $J_d$ in each domain, individuals then evaluate their expected utility from each option $j \in J_d$, denoted by $v_{ij}^d$, by

$$v_{ij}^d = E_{\widetilde{c}}\left[u_i\left(w_i - \lambda_d oop_j(\widetilde{c}) - p_j\right)\right], \tag{2}$$

where expectations are taken over the cost realization $\widetilde{c}$. In addition, $w_i$ is a measure of income or wealth, $oop_j(\widetilde{c})$ captures the out-of-pocket expenditure that is associated with a cost realization of $\widetilde{c}$ under coverage $j$, and $p_j$ denotes the premium associated with coverage option $j$. The parameter $\lambda_d$, which varies across domains but not across individuals, captures context-specific beliefs (or other biases). That is, $\lambda_d = 1$ can be thought of as correct expectations, while $\lambda_d < 1$ ($\lambda_d > 1$) implies biased expectations about risk, which are too optimistic (pessimistic). In the context of the model, $\lambda_d$ enters as biased beliefs, which could be driven by framing effects or probability weighting. More generally, however, one can think of $\lambda_d$ as a "reduced form" way by which we capture a variety of potential domain-specific effects. That is, all else equal, higher (lower) values of $\lambda_d$ require lower (higher) levels of risk aversion to rationalize a given choice, thus providing a free parameter in each domain that captures the level of risk aversion.

To evaluate the expectations for each individual, we abstract from unobservables that may affect ex ante risk, and assume that individuals' risk realization is drawn randomly from the risk realizations of other individuals who are associated with the same group (e.g., based on demographics). That is, if individual $i$ is associated with group $N$ so that $i \in N$, we evaluate individual $i$'s expectations by

$$E_{\widetilde{c}}\left[u_i\left(w_i - \lambda_d oop_j(\widetilde{c}) - p_j\right)\right] = \frac{1}{|N|}\sum_{k \in N} u_i\left(w_i - \lambda_d oop_j(c_k) - p_j\right). \tag{3}$$

Equipped with this model, we can then assume a specific parametric utility function $u_i(\cdot)$ for each individual, such as CARA or CRRA, and map each choice into an interval of coefficients of risk aversion that would rationalize this choice. To see this, note that the set of coverage options in all our domains are vertically ordered (see our discussion in Section 2), so the willingness to pay for incremental coverage is monotone in risk aversion. Conditional on risk expectations, each (discrete) coverage choice can be mapped into an interval of risk aversion parameters that would rationalize the choice. Observing choices of the same individual across different domains, we can now ask whether the intervals associated with these choices overlap. If the answer is positive, it means that there exists a range of domain-general risk aversion coefficients that could generate this individual's choices across the different domains. We can then ask what fraction of individuals have a range of risk aversion coefficients that are consistent across a given set of contexts.

The conceptual approach is similar to the test proposed by Barseghyan et al. (forthcoming), although our use of the $\lambda_d$'s parameters allow us to remain consistent with our "model free" exercise and focus on the consistency of the relative risk preferences of individuals across contexts rather than on the consistency of their absolute levels. This focus likely makes the results less sensitive to modeling assumptions by removing the need to make assumptions (e.g. about the level of risk aversion or the nature of beliefs) as we did in the calibration exercise described above. While our results could speak to the broader question about the consistency of an individual's level of risk aversion across contexts (and, indeed, we mention some such results below), one would naturally worry that in order to infer the level of risk preferences, a richer domain-specific model of risk realization, expectation formation, and coverage choice would be preferred.

### 4.3 Implementation and main results

Using this framework, our empirical exercise attempts to maximize the fraction of the individuals in the sample for whom the implied intervals of risk aversion overlap across two or more domains. We allow the vector of $\lambda_d$'s to be free parameters and search for the set of $\lambda_d$'s that maximize the overlap. Our results have a simple economic interpretation. They represent the fraction of individuals for whom the choices across domains could be rationalized with a single risk aversion parameter, subject to domain-specific effects (that do not vary across individuals). The estimated $\lambda_d$'s (and, in particular, how far they are from 1) can then be interpreted as a measure of how much domain-specific effects is required to rationalize a single risk aversion.

Appendix B provides additional implementation details. To summarize, we assume a CARA utility for the three domains associated with absolute (dollar) risk (health insurance, prescription drug insurance, and dental insurance) and a CRRA utility for the three domains associated with relative (to wage) risk (short- and long-term disability insurance, and 401(k) allocation). We use $\gamma \cdot w_i$ as a multiplicative factor that converts each individual's coefficient of relative risk aversion to absolute risk aversion, where $w_i$ is (in the baseline specification) individual $i$'s observed annual income, and $\gamma$ is an additional free parameter (constant across individuals), which maps annual income to wealth. Other than for this conversion, $w_i$ drops out of the analysis. We search over this additional parameter $\gamma$, in addition to the vector of $\lambda_d$'s, when we search for the maximum overlap.

Table 8 presents the results. Column (2) reports overlap results for all six domains. Column (3) reports overlap results for the five insurance domains. Columns (4) and (5) report results separately for, respectively, the three domains associated with absolute (dollar) risk (health, dental, and drug insurance) and the three domains associated with relative (to income) risk (short- and long-term disability insurance, and 401(k) asset allocation).

Before presenting our baseline results of the maximum fraction of individuals whose implied risk aversion intervals overlap, row 1 presents, as a starting point, the *minimum* fraction of individuals whose implied risk aversion intervals overlap. This can be found by taking the maximum of the fraction of individuals who choose the least risky options and the fraction of individuals who choose the most risky option; with appropriate $\lambda_d$'s, the choices of individuals who always choose the least risky options across domains (or the choices of individuals who always choose the most risky options across domains) can always be rationalized. In our case, the minimum fraction of individuals whose implied risk aversion intervals overlap is given by the fraction of individuals who choose the least risky option in each domain. The first row indicates that, by this metric, at least five percent of the sample can be "mechanically" viewed as consistent across all domains. This number increases substantially, to 26 percent, once we limit the analysis to the five insurance domains.

Row 2 reports the maximum overlap results for our baseline sample. It indicates that, across all six domains, 30 percent of the individuals have implied risk aversion intervals that overlap, once we allow for the domain-specific free parameters (the $\lambda_d$'s). Interestingly, the $\lambda_d$'s required to achieve this overlap are generally well below 1, which is consistent with individuals under-estimating event probabilities. When we only search for overlap across the five insurance domains, we find it to be 38 percent, and it is much higher when we only search for overlap separately across the domains associated with absolute risk (56 percent) and relative risk (70 percent). Naturally, some of this increase in overlap is mechanical, since removing domains (weakly) increases our ability to rationalize the smaller number of choices.

As noted, our introduction of the domain-specific parameter $\lambda_d$ moves the spirit of the analysis away from investigating consistency in an individual's implied *level* of risk aversion across domains toward an analysis of the consistency in an individuals' *ranking* (relative to their peers) of risk aversion across domains. To investigate the importance of these domain-specific "free parameters" for the results, row 3 shows the overlap of risk aversion intervals when we restrict all $\lambda_d$'s (as well as $\gamma$) to be equal to 1. We now find that only 5.3 percent of the sample exhibits choices that overlap in their implied risk aversion intervals. This suggests that the implied levels of risk aversion exhibited may be very different across domains, or that other effects, such as framing or probability weighting, are particularly important in these contexts and different across domains.

Consistent with the "model free" correlation results, our analysis also suggests that the 401(k) domain is the most different. One way to see this is in row 4, where we continue to restrict all five insurance-domains' $\lambda_d$'s (as well as $\gamma$) to be equal to 1, but free up the $\lambda_{401(k)}$ parameter and search for the value that maximizes the overlap. The results illustrate the importance of having a free $\lambda_{401(k)}$ parameter, effectively allowing for very different levels of risk aversion (or beliefs) in this domain. We now obtain a maximum overlap of 28 percent, which is quite different from the overlap of 5 percent when all six $\lambda_d$'s are restricted to be equal to 1, and quite close to the unconstrained maximum overlap of 30 percent (row 2). In other words, allowing a free parameter on $\lambda_{401(k)}$ gets us almost all of the benefit of allowing all six domains-specific free parameters.

### 4.4 Robustness

Somewhat parallel to our robustness analysis in Section 3, we explore the robustness of our model-based results to a number of modeling choices. The results are reported in the subsequent rows of Table 8 (rows 5-13). Overall, the results are reasonably stable across alternative specifications and subsamples.

As noted at the outset, our modeling choices – particularly the introduction of the domain specific free parameters $\lambda_d$'s – was aimed to capture, in a somewhat reduced form way, a wide range of potential domain-specific factors. These include not only domain-specific biases in beliefs or probability weighting functions but other potentially domain specific influences such as the appropriate discount rate, the planning horizon, or framing. As a result, an attractive feature of our modeling approach is that there is a more limited number of domain-specific modeling assumptions with respect to which sensitivity analysis need be evaluated.

However, one domain-specific factor that might contribute to the apparent difference of the 401(k) domain is that the 401(k) choice is continuous, rather than discrete. Reassuringly, row 5 shows that this is unlikely to be important. In particular, we discretize the 401(k) asset allocation decision into three choices (with roughly 40 percent, 16 percent, and 44 percent respectively): contribute nothing to the risk-free funds, contribute everything to the risk-free funds, and contribute in between. For people in this last group, we assign the average contribution share to the risk-free funds among people in this group, which is about 35 percent. The results are indistinguishable from the baseline.

A potentially important modeling choice is the use of CARA utility for three domains and CRRA for the other three, and the free parameter $\gamma$ that is used to convert between them. For this reason, in all rows we have shown results separately for the CARA and CRRA domains (in columns (4) and (5)). We also explored the importance of the free parameter $\gamma$ in our conversion between coefficients of absolute and relative risk aversion. Row 6 shows that constraining this $\gamma$ parameter to be 1 has little effect on the results. Row 7 uses an alternative definition of $w_i$ when we use $\gamma \cdot w_i$ to convert between each individual's coefficient of relative and absolute risk aversion. Specifically, instead of defining $w_i$ as annual income, in row 8 we take account of the individual's 401(k) assets (and the implicit income they generate) by defining $w_i$ as annual income plus five percent of the individual's 401(k) balance. Once again this does not affect the results.

Finally, as noted in our discussion of the "model free" correlation results in the previous section, an important concern with our analysis – particularly for the 401(k) asset allocation decision – is that we do not observe the individual's non-401(k) assets. We therefore subject our model results to the same two types of robustness exercises we performed in Section 3 (see Table 6 in particular). Specifically, we first try to proxy for (and stratify on) housing wealth. For the one third of the sample we were able to obtain housing equity data for, row 8 shows that we estimate a maximum overlap of 29 percent, which is virtually identical to our baseline estimate of 30 percent. This overlap decreases slightly with housing equity (rows 9 to 11); for example, for all domains (column 2) the maximum overlap declines from 32 percent for those with less than \$50,000 in housing equity to 25 percent for those with more than \$150,000 in housing equity. Overall, however, the correlations for strata of individuals with similar housing wealth look very much like the results for the full sample; we interpret these results as suggesting that our estimates are not that sensitive to our lack of data on housing investments. In rows 12 and 13 we compare the overlap across the subsample of employees who do not max out their 401(k) contributions – and therefore are less likely to have outside savings – and the subsample who does. Once again results by strata are very similar to the baseline results.

More generally, across all our various robustness analyses in rows 5 through 13 the maximum fraction of individuals whose implied risk aversion intervals overlap is quite stable, ranging from 25 to 36 percent. The key decision quantitatively appears to be to allow for a domain-specific level of risk aversion in the 401(k) asset allocation (see row 3 and 4); without this the overlap falls considerably.

## 5 Conclusion

This paper investigated the extent to which individuals display a stable ranking in their risk preferences relative to their peers in making market choices over five health-related employer-provided insurance coverage decisions and their 401(k) asset allocation decisions. Our 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 non-trivial magnitudes.

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. The most natural and informative benchmark involved greater modeling assumptions, but the results appear to be quite robust. This in part reflects our strategy of investigating the stability of willingness to take risks relative to one's peers across different domains, rather than the extent to which risk aversion levels are stable across domains. Of course, this choice is not without costs, as it sets a lower hurdle for "domain-generality" of preferences; in a canonical domain-general model of risk aversion, an individual's level of risk aversion would presumably also be constant across contexts.

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 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. The results from our stylized coverage choice model suggest that up to 30 percent of our sample makes choices that may be consistent across all six domains.

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). Results from the stylized coverage choice model also suggest that choices in the 401(k) domain are the most difficult to reconcile with any of the others. 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).

These findings suggest that the extent of domain generality may vary greatly across domains that are more or less "similar" to each other. It would be of great interest in future work to examine the extent of domain generality in more disparate domains than those we currently examine which consisted of five health-related insurance domains and one retirement investment domain. Beyond the data hurdles, however, there is an inherent tension in such an exercise. The more different the domains, the more difficult it is to model and compare consumer choices in a domain-general way. We hope that the approaches outlined here will prove useful in this regard as future work expands to consider a greater set of possibly more disparate domains.

In the meantime, our results may have some implications for current calibration exercises. Calibration work is ubiquitous in the fields of insurance, public finance, and macroeconomics.

The vast majority of this work (including our own past work) attempts to calibrate models using "consensus" parameter estimates (or ranges of estimates) from the literature at large rather than estimates from more similar contexts. The results presented here may suggest that when calibrating models of economic behavior – insurance demand, savings, labor supply, and so on – one might want to consider using preference estimates taken from similar contexts.

## 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 Rutström (2008a). "Lost in state space: are preferences stable?" *International Economic Review*, 49(3), 1091-1112.

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström (2008b). "Eliciting risk and time preferences." *Econometrica*, 76(3), 583-618.

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 (forthcoming). "Are risk preferences stable across contexts? evidence from insurance data." *American Economic Review*, forthcoming.

Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue, and Joshua Teitelbaum (2010). "The nature of risk preferences: evidence from insurance choices." Cornell University, mimeo. Available at http://www.arts.cornell.edu/econ/edo1/BMOT.pdf.

Benjamin, Daniel, Sebastian Brown, and Jesse Shapiro (2006). "Who is 'behavioral'? Cognitive ability and anomalous preferences." Mimeo, Harvard University.

Bucks, Brian K., Arthur B Kennickell, and Kevin B. Moore (2006). "Recent changes in U.S. family finances: evidence from the 2001 and 2004 Survey of Consumer Finances." *Federal Reserve Bulletin*, 92, A1-A38.

Carlin, Caroline, and Robert Town (2010). "Adverse selection: the dog that didn't bite." University of Minnesota, mimeo.

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 microeconometric 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*, 95(2), 238-242.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jurgen Schupp, Gert G. Wagner (2011). "Individual risk attitudes: measurement, determinants and behavioral consequences." *Journal of European Economic Association*, 9(3).

Einav, Liran, Amy Finkelstein, and Mark Cullen (2010). "Estimating welfare in insurance markets using variation in prices." *Quarterly Journal of Economics*, 125(3), 877-921.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf (2010). "Optimal mandates and the welfare cost of asymmetric information: evidence from the U.K. annuity market." *Econometrica*, 78(3), 1031-1092.

Friend, Irwin, and Marshall E. Blume (1975). "The demand for risky assets." *American Economic Review*, 65 (5), 900-922.

Gollier, Christian, and John W. Pratt (1996). "Risk vulnerability and the tempering effect of background risk." *Econometrica*, 64(5), 1109-1123.

Guiso Luigi, Tullio Jappelli, and Daniele Terlizzese (1996). "Income risk, borrowing constraints, and portfolio choice." *American Economic Review,* 86(1), 158-172.

Handel, Benjamin (2010). "Adverse selection and switching costs in health insurance markets: when nudging hurts." UC Berkeley, mimeo. Available at http://www.econ.berkeley.edu/~bhandel/wp/switchingcosts.pdf.

Harrison, Glenn W., John A. List, and Charles Towe (2007). "Naturally occurring preferences and exogenous laboratory experiments: a case study of risk aversion." *American Economic Review*, 75(2), 433-458.

Kimball, Miles S., Claudia R. Sahm, and Matthew D. Shapiro (2009). "Risk preferences in the PSID: individual imputation and family covariation." *American Economic Review*, 99(2), 363-368.

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

Roodman, David. (2009). "Estimating Fully Observed Recursive Mixed-Process Models with cmp". Working Paper Number 168, Center for Global Development. http://www.cgdev.org/content/publications/detail/1421516

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.

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.

Table 1: Employee characteristics in baseline sample

| | Mean | Std. Dev. | 5th pctile | 95th pctile |
|---|---|---|---|---|
| **Panel A: Demographics** | | | | |
| Age | 43.9 | 9.2 | 28 | 58 |
| Annual wage (000$) | 58.4 | 71.7 | 25.6 | 114 |
| Job tenure with Alcoa (years) | 13.2 | 9.6 | 1 | 30 |
| Female | 0.23 | | | |
| White | 0.85 | | | |
| Hourly (non-salary) employee | 0.32 | | | |
| Unionized employee | 0.02 | | | |
| Single coverage tier[a] | 0.19 | | | |
| Number of covered individuals per employee[a] | 2.92 | 1.46 | 1 | 5 |
| **Panel B: Annual Payouts by domain** | | | | |
| Health insurance claims ($) | 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)[b] | 0.061 | | | |
| Long-term disability insurance (fraction with any claims)[c] | 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.

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

[b] Conditional on having a short-term disability claim, the average claim length is 51 days.

[c] 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) |
| **Health Insurance** | | | | |
| 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% | | | |
| **Prescription Drug Insurance** | | | | |
| Option 1 | 23.8% | 181.2 | 248.6 | 385.0 |
| Option 2 | 9.7% | 109.6 | 124.3 | 192.5 |
| Option 3 | 66.4% | | | |
| **Dental Insurance** | | | | |
| Option 1 | 30.0% | 95.7 | 45.2 | 112.9 |
| Option 2 | 70.0% | | | |
| **Short-Term Disability Insurance**[a] | | | | |
| Option 1 | 15.5% | 165.1 | 140.2 | 825.7 |
| Option 2 | 17.9% | 63.5 | 70.3 | 413.4 |
| Option 3 | 66.6% | | | |
| **Long-Term Disability Insurance**[a] | | | | |
| Option 1 | 16.3% | 152.4 | 17.0 | 395.7 |
| Option 2 | 14.9% | 63.5 | 8.5 | 197.9 |
| Option 3 | 68.8% | | | |
| **401(k) allocation**[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.

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

[b] 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(a): Correlation estimates, without controls

**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: Correlation estimates from a system of ordered probits**

| | Health | Drug | Dental | STD | LTD |
|---|---|---|---|---|---|
| Drug | 0.550 | | | | |
| Dental | 0.339 | 0.410 | | | |
| STD | 0.292 | 0.303 | 0.271 | | |
| LTD | 0.243 | 0.298 | 0.266 | 0.768 | |
| 401(k) | 0.055 | 0.071 | 0.046 | 0.032 | 0.020 |
| | | | | (0.004) | (0.069) |

Average correlation is 0.264

**Panel C: Correlation estimates from a 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

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. Panel B shows results from a system of five ordered probits and one linear regression for the 401(k) domain (see text for more details). Panel C reports the correlation structure from the multivariate regression shown in equation (1). Both Panel B and Panel C include control (indicator) variables for the benefit menu the employee faces; for Panel B, we exclude all menus that were offered to fewer than 100 people, reducing the sample size by 86 employees.

Table 3(b): Correlation estimates, controlling for predictors of risks

**Panel A: Correlation estimates from a system of ordered probits**

| | Health | Drug | Dental | STD | LTD |
|---|---|---|---|---|---|
| Drug | 0.494 | | | | |
| Dental | 0.302 | 0.409 | | | |
| STD | 0.249 | 0.245 | 0.258 | | |
| LTD | 0.210 | 0.250 | 0.255 | 0.764 | |
| 401(k) | 0.036 | 0.043 | 0.037 | -0.005 | -0.006 |
| | (0.001) | | (0.003) | (0.644) | (0.562) |

Average correlation is 0.234

**Panel B: Correlation estimates from a multivariate regression**

| | Health | Drug | Dental | STD | LTD |
|---|---|---|---|---|---|
| Drug | 0.411 | | | | |
| Dental | 0.208 | 0.250 | | | |
| STD | 0.155 | 0.156 | 0.156 | | |
| LTD | 0.130 | 0.157 | 0.153 | 0.593 | |
| 401(k) | 0.038 | 0.032 | 0.026 | 0.002 | -0.002 |
| | | | (0.004) | (0.859) | (0.825) |

Average correlation is 0.164

The reports results for our baseline sample of 12,752 employees. Panels A and B are analogous to Panels B and C in Table 3(a), respectively. The results reported in this table include additional eleven control variables for predicted and realized risk in each equation. These attempt to control for heterogeneous risk expectations across individuals, which may be correlated across domains. See the text (Section 3.1) for additional details. As in Table 3(a), both panels include also control (indicator) variables for the benefit menu the employee faces; for Panel A, we exclude all menus that were offered to fewer than 100 people, reducing the sample size by 86 employees.

Table 4(a): Summary correlations by groups, ordered probit specification

| | | Obs. (1) | Average correlation (2) | Health-Drug correlation (3) | Health-STD correlation (4) | Health-401(k) correlation (5) |
|---|---|---|---|---|---|---|
| (1) | Single coverage | 2420 | 0.309 | 0.643 | 0.379 | 0.082 |
| | Non single | 10246 | 0.251 | 0.517 | 0.267 | 0.052 |
| (2) | More tenured | 11641 | 0.262 | 0.547 | 0.287 | 0.058 |
| | Newly hired | 1025 | 0.269 | 0.569 | 0.289 | 0.012 |
| (3) | Higher wage | 3145 | 0.240 | 0.524 | 0.198 | 0.078 |
| | Lower wage | 3126 | 0.246 | 0.534 | 0.336 | 0.029 |
| (4) | Don't allocate to Alcoa Stock | 7241 | 0.272 | 0.548 | 0.300 | 0.066 |
| | Allocate to Alcoa stock | 5245 | 0.252 | 0.552 | 0.277 | 0.036 |
| (5) | Rebalance 401(k) portfolio | 3610 | 0.261 | 0.551 | 0.264 | 0.080 |
| | Don't rebalance | 9056 | 0.266 | 0.551 | 0.302 | 0.047 |
| (6) | Over 55 years old | 1690 | 0.248 | 0.595 | 0.251 | 0.062 |
| | Under 35 years old | 2550 | 0.276 | 0.539 | 0.326 | 0.032 |
| (7) | Salaried employees | 8594 | 0.256 | 0.541 | 0.256 | 0.068 |
| | Hourly employees | 4072 | 0.247 | 0.542 | 0.326 | 0.014 |

The table reports the correlation coefficients for the subsamples specified in the row headers. The estimates all use Panel B of Table 3(a) as a baseline. That is, we report the correlation structure of the residuals from estimating the system of ordered probit equations (with a single linear equation for 401(k) choice), 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 (at least once) and did not rebalance their 401(k) portfolio during the year.

Table 4(b): Summary correlations by groups, multivariate regression

| | | Obs. (1) | Average correlation (2) | Health-Drug correlation (3) | Health-STD correlation (4) | Health-401(k) correlation (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 fully parallels Table 4(a), except that it uses the residuals from estimating the multivariate regression specification (Panel C of Table 3(a)), as shown in equation (1), as a baseline.

Table 5: Robustness I – Alternative specifications and samples definitions

| | | Obs. (1) | Average correlation (2) | Health-Drug correlation (3) | Health-STD correlation (4) | Health-401(k) correlation (5) |
|---|---|---|---|---|---|---|
| **Panel A. A system of ordered probits** | | | | | | |
| 1 | Baseline specification | 12,666 | 0.264 | 0.55 | 0.292 | 0.055 |
| 2 | Discretizing the 401(k) choice | 12,666 | 0.260 | 0.550 | 0.292 | 0.050 |
| 3 | Control for coverage tier | 12,666 | 0.264 | 0.546 | 0.292 | 0.056 |
| 4 | Use only the largest pricing menu | 7,722 | 0.268 | 0.552 | 0.277 | 0.067 |
| 5 | Include those in opt-out and HMO | 15,399 | 0.230[a] | -- | -- | -- |
| 6 | Include employees who did not contribute to 401(k) | 15,344 | 0.368[b] | 0.540 | 0.295 | -- |
| 7 | Include those not offered LTD coverage | 15,570 | 0.230[c] | 0.540 | 0.292 | 0.052 |
| 8 | Exclude those in Health Option 1 (due to HRA component) | 10,473 | 0.223 | 0.317 | 0.280 | 0.009 |
| 9 | Include only new hires | 1,025 | 0.269 | 0.569 | 0.289 | 0.012 |
| 10 | Exclude individuals who may have chosen default options | 11,243 | 0.279 | 0.627 | 0.328 | 0.067 |
| **Panel B. Multivariate regressions** | | | | | | |
| 1 | Baseline specification | 12,752 | 0.188 | 0.452 | 0.188 | 0.057 |
| 2 | Discretizing the 401(k) choice | 12,752 | 0.184 | 0.452 | 0.188 | 0.045 |
| 3 | Control for coverage tier | 12,752 | 0.186 | 0.447 | 0.187 | 0.058 |
| 4 | Use only the largest pricing menu | 7,722 | 0.195 | 0.452 | 0.191 | 0.069 |
| 5 | Include those in opt-out and HMO | 15,409 | 0.165[a] | -- | -- | -- |
| 6 | Include employees who did not contribute to 401(k) | 15,402 | 0.257[b] | 0.446 | 0.184 | -- |
| 7 | Include those not offered LTD coverage | 15,675 | 0.162[c] | 0.442 | 0.183 | 0.052 |
| 8 | Exclude those in Health Option 1 (due to HRA component) | 10,547 | 0.147 | 0.226 | 0.175 | 0.009 |
| 9 | Include only new hires | 1,044 | 0.195 | 0.472 | 0.184 | 0.023 |
| 10 | 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. Analogously to Table 3(a), Panels B and C respectively, Panel A uses the system of ordered probits and Panel B uses multivariate regressions. Column (2) shows the simple average of the 15 pairwise correlations, and columns (3) through (5) report correlations for specific pairs. For ease of comparison, row 1 replicates the baseline specification from Table 3(a). Each row shows a single deviation from the baseline specification. Row 2 replaces the continuous 401(k) measure with a discretized ordinal measure ranging from 1 to 3, row 3 includes coverage tier (based on health coverage) fixed effects, and row 4 reports results using the largest (modal) benefit menu (and therefore does not require menu fixed effects). Rows 5-10 report results from alternative samples. In rows 5, 6, and 7 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 9 we limit the sample to new hires (defined as job tenure at Alcoa of less than two years). In row 10 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.

[a] The comparable average correlation (that is, over the 6 pairs that do not include health and drug coverage) in the baseline specification is 0.234 (Panel A) and 0.169 (Panel B).

[b] The analogous average correlation (that is, over the 10 pairs that do not include 401(k) choices) in the baseline specification is 0.374 (Panel A) and 0.262 (Panel B).

[c] The analogous average correlation (that is, over the 10 pairs that do not include long-term disability coverage) in the baseline specification is 0.237 (Panel A) and 0.169 (Panel B).

Table 6: Robustness II – Non-Alcoa investments

| | | Obs. (1) | Average correlation (2) | Health-Drug correlation (3) | Health-STD correlation (4) | Health-401(k) correlation (5) | Drug-401(k) correlation (6) | Dental-401(k) correlation (7) | STD-401(k) correlation (8) | LTD-401(k) correlation (9) |
|---|---|---|---|---|---|---|---|---|---|---|
| **Panel A. A system of ordered probits** | | | | | | | | | | |
| 1 | Baseline specification | 12,666 | 0.264 | 0.55 | 0.292 | 0.055 | 0.071 | 0.046 | 0.032 | 0.02 |
| 2 | Housing Subsample | 4,278 | 0.271 | 0.541 | 0.298 | 0.049 | 0.088 | 0.055 | 0.009 | 0.007 |
| 3 | House Equity < $50,000 | 1,362 | 0.282 | 0.502 | 0.343 | 0.027 | 0.087 | 0.091 | 0.005 | 0.015 |
| 4 | Housing Equity $50,000-$150,000 | 1,523 | 0.282 | 0.592 | 0.306 | 0.058 | 0.074 | 0.001 | 0.018 | 0.014 |
| 5 | Housing Equity > $150,000 | 1,355 | 0.253 | 0.514 | 0.239 | 0.065 | 0.104 | 0.081 | 0.019 | -0.004 |
| 6 | Maxed out 401(k) contributions | 1,731 | 0.288 | 0.608 | 0.305 | 0.114 | 0.071 | 0.036 | 0.011 | 0.004 |
| 7 | Did not max out 401(k) contributions | 10,935 | 0.258 | 0.539 | 0.285 | 0.044 | 0.070 | 0.046 | 0.032 | 0.023 |
| **Panel B. Multivariate regressions** | | | | | | | | | | |
| 1 | Baseline specification | 12,752 | 0.188 | 0.452 | 0.188 | 0.057 | 0.056 | 0.035 | 0.029 | 0.018 |
| 2 | Housing Subsample | 4,309 | 0.195 | 0.441 | 0.203 | 0.051 | 0.070 | 0.041 | 0.016 | 0.012 |
| 3 | House Equity < $50,000 | 1,399 | 0.202 | 0.410 | 0.229 | 0.042 | 0.075 | 0.072 | 0.016 | 0.017 |
| 4 | Housing Equity $50,000-$150,000 | 1,544 | 0.199 | 0.488 | 0.211 | 0.049 | 0.052 | 0.000 | 0.013 | 0.014 |
| 5 | Housing Equity > $150,000 | 1,366 | 0.184 | 0.417 | 0.167 | 0.066 | 0.083 | 0.061 | 0.022 | 0.003 |
| 6 | Maxed out 401(k) contributions | 1,740 | 0.212 | 0.499 | 0.225 | 0.089 | 0.049 | 0.030 | 0.016 | 0.004 |
| 7 | Did not max out 401(k) contributions | 11,012 | 0.181 | 0.441 | 0.176 | 0.050 | 0.056 | 0.034 | 0.029 | 0.020 |

This table reports correlation results for various subsamples. Analogously to Table 3(a), Panels B and C respectively, Panel A uses the system of ordered probits and Panel B uses multivariate regressions. Column (2) shows the simple average of the 15 pairwise correlations, and columns (3) through (9) report correlations for specific pairs. For ease of comparison, row 1 replicates the baseline specification from Table 3(a). Row 2 presents the results for approximately one-third of the sample for which we were able to match data on their housing equity. Rows 3 through 5 present results for various subsamples of this "housing subsample," as indicated. Rows 6 and 7 present results separately for individuals who have maxed out their possible 401(k) contributions and those who have not.

Table 7: 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 | 0.070 | 0.107 | 0.056 | 0.043 | 0.023 | 0.024 |
| 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 | 0.247 | 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 benefit 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 Section 3.1 for more details). "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 8: Model-based results

| | | Obs. (1) | All domains (2) | All insurance domains (3) | Three CARA domains (4) | Three CRRA domains (5) |
|---|---|---|---|---|---|---|
| 1 | Minimum overlap | 11,898 | 5% | 26% | 35% | 10% |
| 2 | Baseline specification | 11,898 | 30% | 38% | 56% | 70% |
| 3 | Restricted: $\lambda = \gamma = 1$ | 11,898 | 5% | 31% | 44% | 14% |
| 4 | Restricted: flexible $\lambda$ only on 401(k); $\gamma = 1$ | 11,898 | 28% | -- | -- | 61% |
| **Alternative specifications:** | | | | | | |
| 5 | Discretize 401(k) | 11,898 | 30% | -- | -- | 70% |
| 6 | Restricted: $\gamma = 1$ | 11,898 | 30% | 36% | -- | -- |
| 7 | Alternative definition of income | 11,898 | 30% | 38% | -- | -- |
| **Alternative samples:** | | | | | | |
| 8 | Housing Subsample | 4,054 | 29% | 35% | 56% | 68% |
| 9 | House Equity < \$50,000 | 1,305 | 32% | 38% | 57% | 70% |
| 10 | Housing Equity \$50,000-\$150,000 | 1,453 | 30% | 36% | 57% | 69% |
| 11 | Housing Equity > \$150,000 | 1,296 | 25% | 32% | 54% | 64% |
| 12 | Maxed out 401k contributions | 9,394 | 29% | 35% | 55% | 69% |
| 13 | Did not max out 401k contributions | 2,504 | 36% | 44% | 62% | 71% |

This table reports results from the exercise described in Section 4 (and in additional detail in Appendix B). Each entry in columns (2) through (5) represents our estimate of the fraction of individuals whose entire vector of choices (as given by the column header) could be rationalized given the analogous specification (as given by the row header). Specifically, column (2) reports the fraction of individuals whose estimated ranges of risk aversion in each domain overlap across all six domains; column (3) reports the fraction with overlap across the five insurance domains (that is, not including 401(k) allocation), column (4) reports the fraction with overlap across the three domains associated with absolute risk (health, drug, and dental), and column (5) reports the fraction with overlap across the three relative risk domains (short- and long-term disability, and 401(k) allocation).

Each row reports a different specification. The first row report the minimum fraction of individuals with overlap in their risk aversion ranges; this is the fraction of individuals who always choose the least risky option in each domain. The second row reports our baseline specification, as described in the text. All other rows report variants of the baseline, each with a single deviation from the baseline as described. In row 3 we constrain all six $\lambda_d$'s and $\gamma$ to be 1. In row 4 we restrict $\gamma$ and five of the $\lambda_d$'s to be 1 but free up the $\lambda_{401(k)}$ parameter. In row 5 we discretize the 401(k) asset allocation decision into three choices: invest nothing in the risk-free asset, invest all in the risk-free asset, or "in between", which we parameterize based on the average risk free share (35 percent) of those in this category. In row 6 we restrict $\gamma$ to be 1. In row 7 we define income (used to convert between each individual's coefficient of relative and absolute risk aversion) as annual income plus 5 percent of 401(k) balances, instead of as annual income as in the baseline specification. Rows 8 through 13 reports results from the baseline specification using various subsamples of our population. Specifically, in row 8 we limit the results to the sample for whom we were able to link in data on housing wealth. Rows 9 through 11 show results stratified by housing equity level. Rows 12 and 13 split the sample between those who have contributed the maximum possible amount to their 401(k) and those who have not maxed out their possible 401(k) contributions.

# NOT FOR PUBLICATION APPENDICES

**Appendix A: A calibrated model of short-term and long-term disability insurance choices.**

In this appendix we describe the details of the calibration exercise on which we report in the beginning of Section 4. The objective of the calibration exercise is to illustrate how one could produce 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. This allows us to rely on a single choice model for both domains, rather than on a domain-specific model. 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.

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

$$Eu(c) = E_{\tilde{d}}\left[\left(1-\tilde{d}+RR_j*\widetilde{d}-p_j\right)^{1-\gamma}\right], \tag{4}$$

where $\widetilde{d}$ is the (ex-ante random) fraction of days in a year the employee claims (due to disability), $RR_j$ is the wage replacement rate associated with coverage $j$, the premium $p_j$ 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 an annual discount factor of .95 for long-term disability.

We assume that the distribution of $\gamma$ in the population to be lognormal with parameters $\mu$ and $\sigma$, such that the values of $\mu$ and $\sigma$ are chosen to produce an average relative risk aversion coefficient of 3 or 0.7 (depending on the specification) and a coefficient of variation of risk aversion of 10. The coefficient of variation (of 10) matches the estimates reported by Cohen and Einav (2007). Cohen and Einav (2007) mention higher numbers of relative risk aversion, but they essentially estimate absolute risk aversion, so mapping it to this lower levels of relative risk aversion amounts to simply assuming lower relevant wealth (the simulated correlations remain about the same when we instead use an average coefficient of risk aversion of 30, and maintain the same coefficient of variation). 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.

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.

## Appendix B: Implementation details for the model-based approach of Section 4

This appendix provides additional details that underlie the baseline exercise reported in Section 4 (and Table 8).

*Health, Dental, and Drug coverage.* The risks in these three domains are measured in dollars. Therefore, for our baseline estimates, we assume a CARA utility function in these domains. That is, we use equation (3) to compute individual $i$'s expected utility from option $j$ by substituting $u_i(x) = -\exp(-r_i x)$, incorporating the plan details (regarding deductible and out-of-pocket maximum) to compute $oop_j(c)$, and grouping individuals into risk categories by their coverage tier (single coverage, employee plus spouse, employee plus children, and family coverage) and randomly drawing from individuals' realization of total medical expenditure $c$.

*Short-term and Long-term disability coverage.* The risks (and premiums) in these domains are all proportional to the employee's (annual) wage. It is therefore natural to assume a CRRA utility function for these two domains. Again, we use equation (3) to compute individual $i$'s expected utility from option $j$ by substituting $u_i(x) = x^{1-\gamma_i}$, assuming all individuals are grouped at the same risk, drawing the claimed disability days for each individual, and computing $oop_j(d) = (1 - RR_j)d$ where $RR_j$ is the wage replacement rate associated with coverage $j$. A minor complication arises in the case of long-term disability coverage, where the data on realized risks is slightly censored (for long spells of disability absence), so we impute the full predicted absence based on the observed propensity to remain on (long-term) disability between the first and second year.

*Determining cutoffs and defining intervals.* Given a value of $\lambda_d$ for domain $d$ with $J$ options, we partition $[0, \infty]$ into $J$ intervals $[r_1(\lambda_d) = 0, r_2(\lambda_d)]$, ..., $[r_k(\lambda_d), r_{k+1}(\lambda_d)]$, ..., $[r_J(\lambda_d), r_{J+1}(\lambda_d) = \infty]$, such that an individual with a given distribution of risk and a risk aversion parameter in interval $[r_k(\lambda_d), r_{k+1}(\lambda_d)]$ will choose option $k$. For a given $\lambda_d$, a menu of options and distribution of risk, we first find the level of risk aversion $r_{k,k+1}(\lambda_d)$ such that an individual is indifferent between choices $k$ and $k+1$, where option $k+1$ has the higher premium and higher coverage. There are a couple of cases to bear in mind:

- If a risk neutral individual prefers option $k+1$ over option $k$ then option $k$ is dominated and choice $k+1$ cannot be rationalized. In such a case, some of the intervals will be empty.
- For lower values of lambda, the risk might be small enough that option $k$ should never be chosen. In the limit, if $\lambda = 0$, then only the lowest coverage option can be rationalized, and again some of the intervals are empty.
- For all other cases, we can find a cutoff value such that an individual faced with option $k$ and $k+1$ will choose option $k$ for $r < r_{k,k+1}$, and option $k+1$ if $r > r_{k,k+1}$.

Using the procedure described above, we compute $J(J-1)/2$ cutoff values for each pair of options, which define the relevant intervals of risk aversion implied by each coverage choice in the data.

*401(k) choice.* Here, because the decision is continuous, we slightly deviate from the description provided in the paper, and instead rely on the same exercise performed in the seminal paper of Friend and Blume

(1975). As they show, one can convert one's share invested in a risky asset $\alpha_i$ to one's coefficient of relative risk aversion $\gamma_i$, by applying $\gamma_i = \frac{1}{\alpha_i}\frac{E(r_m - r_f)}{\sigma_m^2}$. Our inclusion of a domain-specific adjustment $\lambda_{401(k)}$ simply implies that $\lambda_{401(k)}$ multiplies the right-hand-side, illustrating how this manipulation frees up the level of risk aversion. We use the average return of the safe funds to compute the (monthly) risk free return $E(r_f) = 0.0036$. We aggregate all the funds in our sample invested in the risky funds to compute an estimate of the expected (monthly) return of the "risky" asset and its standard deviation, which are given by $E(r_m) = 0.0103$ and $\sigma_m = 0.0285$. Taken together, this implies that $\frac{E(r_m - r_f)}{\sigma_m^2} = 8.35$. We further assume that people who invest all their 401(k) contributions in the risky funds are at a corner solution, implying that for such individuals we obtain that $\gamma_i \in \left[0, \lambda_{401(k)}\frac{E(r_m - r_f)}{\sigma_m^2}\right]$.

*Conversion between absolute and relative risk aversion.* For each individual we have three intervals for their value of absolute risk aversion, based on their health, dental, and drug coverage choice, and three intervals (or a point in some cases for 401(k)) for the value of their relative risk aversion from short-term and long-term disability coverage and their 401(k) allocation. To evaluate the consistency of choices across all six domains, we need to convert the absolute risk aversion intervals to relative-risk aversion. We use another free parameter $\gamma$ (which could be interpreted as converting annual income to wealth), as well as the data on the individual's annual income, such that $RRA = ARA \times wage \times \gamma$, where $RRA$ and $ARA$ are the coefficients of relative and absolute risk aversion, respectively.

Appendix Table A1: Coverage Details for Insurance Plans

| | Summary of Key Coverage Details (1) | Additional details (2) |
|---|---|---|
| Health Insurance[a] | Deductible (In-network / out-of-network) | |
| Option 1[b] | 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[c] | Wage replacement rate | |
| Option 1 | mostly 60% (sometimes 40%) | |
| Option 2 | mostly 80% (sometimes 60%) | 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 3 | mostly 100% (sometimes 80%) | |
| Long-Term Disability Insurance[c] | Replacement rate | |
| Option 1 | mostly 50% | |
| Option 2 | mostly 60% | All long-term disability coverage is payable after 26 weeks of disability (when the shirt-term disability coverage is capped). |
| 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.

[a] Health insurance: deductibles are shown for the non-single coverage tier; deductibles for single coverage are half what is shown.

[b] Option 1 includes a Health Reimbursement Account (HRA) in which Alcoa contributes $1,250 in tax free money each year that the employee can used 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.

[c] Short-term and Long-term disability benefits (column (1)) are proportional to the employee's wage.

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

| Fund name (Asset Class) | | Monthly return | | | |
|---|---|---|---|---|---|
| | Share[a] | Mean | Std. Dev. | Min. | Max. |
| Classified (by us) as "Risk Free": | | | | | |
| 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 |
| All other classified as risky: | | | | | |
| 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 |
| Benchmarks during the same period: | | | | | |
| Risk free[b] | -- | 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).

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

[b] 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.