

**Complex Decision Making: The Roles of Cognitive Limitations,
Cognitive Decline and Ageing**

by

Michael P. Keane

University of Oxford and University of New South Wales*

Susan Thorp

The University of Sydney[^]

September 7, 2015

Revised: September 9, 2016

* Rm 2111 (Phase II, 2nd Floor), Manor Road Building, Oxford OX1 3UQ, U.K.; T: 44 (0)1865 28148; Email: michael.keane@nuffield.ox.ac.uk

[^]Rm 539, Level 5, Codrington Building H69, The University of Sydney, NSW 2006 Australia; T: 61 (0)2 90366354; Email: susan.thorp@sydney.edu.au

Acknowledgements: This chapter has been prepared for *The Handbook of Population Ageing*, Elsevier, J. Piggott and A. Woodland (eds.). Keane's work on this project has been supported by Australian Research Council grants FF0561843 and FL110100247. Thorp's work on this project has been supported by Australian Research Council grant DP120102239. We thank the editors and two anonymous referees for exceptionally useful comments.

Complex Decision Making: The Roles of Cognitive Limitations, Cognitive Decline and Ageing

Abstract:

We review evidence on decision making in complex choice situations – i.e., situations where there are many alternatives and/or where attributes of alternatives are difficult to understand. We focus on choices about health insurance, health care, and retirement planning, all of which are very important for the well-being of the elderly. Our review suggests that consumers in general, and the elderly in particular, have great difficulty making optimal choices in these areas. They often behave in ways that imply a high degree of “confusion,” such as (i) failure to understand key attributes of alternatives, or (ii) inadequate cognitive capacity to process payoff relevant information. We go on to discuss extensions to standard rational choice models that account for consumer confusion. These include allowing perceived attributes to depart from true attributes; the use of heuristics; and inattention or procrastination. Such departures from rationality can be moderated by cognitive ability, age etc. We hope that these new models may be useful in designing paternalistic interventions.

Keywords: Aging; Life cycle; Health insurance; Health care; Pensions; Retirement plans;

Discrete choice models

JEL codes: I13; I11; J14; J32; H55; D14; D83; D84; D91; C35

1. Introduction

Over the past 30 years, many public policy experts have advocated greater consumer choice in areas that are of particular relevance to the elderly. These areas include health care, health insurance and retirement planning. The rationale for greater choice is that it should generate more competition among service providers. This, in turn, should lead to lower prices and higher quality services. Governments in the U.S., U.K. and Australia have all adopted such policies (albeit to a greater or lesser extent).

The foundation of the “more choice is *necessarily* good” argument rests on the rational choice paradigm. This assumes consumers have both adequate information and adequate mental capacity to understand the choices they face. For example, as Frank (2004) notes: “There is a presumption in much of health economics that more choice is better... the *de facto* model of health care delivery in the U.S. and other nations is that of “managed competition” (Enthoven, 1988). The assumption is that consumers find the right health plans and that overall the net gains of wider choice are positive.”¹

But there is considerable debate about whether people in general, or the elderly in particular, can actually understand the complex choices they face in markets such as health care, health insurance, retirement benefits and long-term care. This chapter will examine the empirical evidence on this topic, and try to assess whether the “more choice is always good” argument is tenable for such complex products. If many people have difficulty making rational decisions in these areas, all of which are crucial for the well-being of the older population, then this is a serious source of concern with respect to population ageing.

Before proceeding, it is important to clarify what we mean by a “complex” choice. A good starting point may be to establish what we would call “simple” choices. For instance, in choosing a laundry detergent, one could reasonably argue there are only four main attributes

¹ Of course, even if consumers have mental or informational limitations, having more choices may often be utility enhancing. But it becomes an empirical question, and the answer is context dependent.

(cleansing power, scent, softening and price), and so a consumer's task in finding a laundry detergent he/she likes is rather simple. It is made even simpler by the fact that the consumer can experiment with different detergents at low cost – see Erdem and Keane (1996).

In contrast, a “complex” choice situation may arise in two main ways:

- a) The object under consideration is complex, in that it has many attributes, or some attributes that are difficult to understand or evaluate;
- b) The choice set is complex because there are a very large number of alternatives.

A good example of a complex choice object is a health insurance plan. These typically consist of a complex set of state contingent payout rules, covering many different health conditions, types of treatment, and types of providers.² Thus, for any particular plan, it is difficult to determine one's expected out-of-pocket health care costs. The quality of care that a plan provides is also very hard to measure. Likewise, superannuation or pension plans are also very complex choice objects, as they consist of multiple contribution, investment, insurance and advice structures, along with various decumulation strategies.

Large choice sets generate complex choice situations in two main ways. First, even if the entire choice set can be readily observed, it may be impractical for consumers to consider all alternatives (as standard choice models assume they do).³ Second, it may be difficult (or costly) to even discover all the available options.

In this chapter we discuss evidence on how people behave in complex choice situations, with special reference to choices in the areas of health insurance, health care, and retirement planning. We focus on these topics because the well-being of senior citizens depends critically on people making “good” choices in these areas, not just in old age but

² For example, what happens if a resident of Minnesota catches valley fever in Phoenix and has to be treated at an out-of-network emergency clinic?

³ Consider standing in the cereal aisle of the supermarket. The entire choice set of roughly 100 options is clearly visible, but no one would have time to carefully search through it all – see Keane and Wasi (2012). A closely related point is that, even given complete information about all alternatives, a large choice set makes comparison of attributes across alternatives difficult.

over the whole life-cycle.

An important point is that the rational choice paradigm assumes all economic agents are capable of making choices that maximize expected utility. However, it is intuitive that, when confronted with complex choices, some consumers will make “better” decisions than others. Some people may be better able to handle complex choice situations for a number of reasons: higher cognitive ability, more patience, better decision making skills, access to assistance, etc.. A striking implication is that increased consumer sovereignty may have distributional implications, favouring those who are better decision makers.

Of course, people of all ages may have difficulty in complex decision making environments. But the problem of cognitive decline with age may well mean that senior citizens will have greater difficulties in an environment of enhanced consumer sovereignty. In particular, a person’s fluid abilities, their capacity to solve problems and think logically in novel situations, declines from early adulthood. However age-related reductions in fluid abilities are offset to some degree by growth in stored knowledge, accumulated experience and mastery of life – known as “crystallized abilities” (Blazer et al., 2015).

Attitudes to risk also change with age, with older adults less interested in sensation-seeking, more risk avoidant, but no less patient, than younger adults (Roalf et al., 2011). Thus, aging may not only affects one’s capacity to make complex choices, but also the preferences and perceptions of well-being that determine choices. Older individuals also show a positivity bias, or socio-emotional selectivity. This bias directs memory and attention to positive emotions and information. While selectively “pruning” negative experiences can promote feelings of well-being, it can also make the elderly susceptible to fraud and manipulations of trust (Castle et al., 2012; Mather and Carstensen 2005). This can obviously create problems in the domain of financial decision making.

The effects of diminishing skills at later ages can be lessened if people recognize

these changes and get help, but many, understandably, put this off, for fear of losing their independence (Blazer et al., 2015). What's more, as people experience decreasing financial ability at older ages, their confidence in their financial skill does not appear to drop off commensurately (see, e.g., Hanoch et al., 2009; Gamble et al., 2014a). Thus, many who should get help don't (Gamble et al., 2014a).

There is evidence that people can compensate for age-related declines in some abilities using cognitive reserves (Stern 2002). Reserves are partly genetically determined but also built up by enriching mental stimulation and physical activities (Hertzog et al., 2008). However, retirement itself seems to exacerbate cognitive decline, most likely because of lower stimulation and less incentive to maintain human capital by building cognitive reserves (Mazzonna and Peracchi 2012; Grotz et al., 2015; Bonsang et al., 2011). For all these reasons, it is particularly interesting to analyse the behaviour of senior citizens in complex choice environments.

The outline of the chapter is as follows: In Section 2 we examine evidence of consumer choice in the areas of health insurance and health care. We emphasize how consumers exhibit symptoms of “confusion” when choosing among insurance plans, in the sense that they make choices in ways that suggest they do not understand the attributes of insurance plans very well. We return to the theme of “confusion” in choice behavior throughout the chapter. In Section 3 we focus on decisions related to retirement planning. In Section 4 we discuss ideas on how to model “choice under confusion.” Section 5 concludes.

2. Health Insurance and Health Care Choices

Here we survey the evidence on how people in general, and senior citizens in particular, act when confronted with complex choices about health insurance and health care. We also look at the evidence on how attempts to simplify the choice environment (e.g., product standardization) affect behavior at the individual and aggregate level.

2.1. Evidence of “Confusion” in Making Health Insurance Choices

There is considerable evidence that people have problems making judgments involving probability and risk, which means they generally have problems making good choices about insurance or investment products (see, e.g., Johnson et al., (1993), Peters, Hibbard et al., (2007) and Peters (2008)). And there is also evidence that these difficulties increase with age (see, e.g., Peters, Hess, et al., 2007; Cole et al., 2008; Samanez-Larkin et al., 2010).

Turning specifically to health insurance, quite a few papers have appeared in the econometric literature finding evidence of confusion in buying private health insurance. Early studies of this type were Harris and Keane (1998), McFadden (2006), Winter et al., (2006), Fang et al., (2008); Abaluck and Gruber (2009), Maestas et al., (2009) and Frank and Zeckhauser (2009). Many other papers have followed.

2.1.1. Evidence that Consumers Fail to Understand Health Plan Attributes

Harris and Keane (1998) found that senior citizens have fundamental misperceptions about key attributes of their health insurance options. It is worth describing their work in some detail, not only because it is one of the early papers finding econometric evidence of “confusion,” but, more importantly, because their method for detecting “confusion” – by estimating choice models that allow for divergence between perceived and true attributes of alternatives – may be a useful way to relax the rational choice paradigm in other contexts.

To proceed, Harris and Keane (1998) – henceforth HK – modeled the health insurance choices of senior citizens living in Minneapolis and St. Paul, Minnesota, using data collected by the Health Care Financing Administration (HCFA) in 1988. To understand the choice problem faced by these consumers, it is important to understand two things about this market.

First, all senior citizens in the US have federally funded health insurance under Medicare. However, the basic Medicare “fee-for-service” program requires significant cost

sharing in the form of deductibles and co-pays, and leaves a number of services, such as preventive care and, until recently, prescription drugs, uncovered.⁴ Thus, many senior citizens buy supplemental insurance, known as “Medigap,” that cover these “gaps” in Medicare. There were many Medigap plans offered by private insurers in Minneapolis/St. Paul in 1988. But, as Fang et al., (2008) note, plan features are highly regulated. Thus, HK found that all plans could be fairly accurately categorized into just two types: those with and without drug coverage, with other plan features (like premiums) fairly comparable within each type.

Second, as an alternative to supplemental insurance, seniors can also join a managed care plan offered by a private firm. Basically, a managed care plan offers more complete coverage than basic Medicare, but at the cost of restricting provider choice or otherwise constraining consumer behavior.⁵ There are two basic types of managed care plan, known as independent practice associations (IPA) and group health maintenance organizations (HMO). In an IPA, a private insurer contracts with a set of health care providers, and plan members can choose to obtain services from any provider in the network.⁶ In a group HMO, the private insurer employs a staff of providers, and provider choice is sharply curtailed.⁷

Thus, the choice set contained five options: (i) Basic Medicare, (ii) Medicare plus a Medigap plan without drug coverage, (iii) Medicare plus a Medigap plan with drug coverage, (iv) an IPA, or (v) a Group HMO. Key attributes of plans are described in Table 1. These are: the premium; drug coverage; preventive care; provider choice; and whether an enrollee must submit claims for reimbursement after using medical services.

⁴ The Medicare Modernization Act of 2004 introduced partial drug coverage. The new benefit did not take effect until 2006, and it still left substantial cost sharing requirements.

⁵ Medicare HMOs receive a per enrollee government payment that is less than the government’s cost of insuring a typical Medicare enrollee. If the HMO serves the person for less than the amount of the subsidy, it makes a profit and the government saves money. Of course, the arrangement is problematic if the HMO saves on costs via cherry picking its enrollees rather than through enhanced efficiency.

⁶ The idea is that the IPA can obtain cost savings by negotiating favorable reimbursement rates with the providers who join. Ideally then, these providers have to contain costs in order to still make profits from serving the IPA patients, so efficiency of health care provision is enhanced.

⁷ A group HMO combines the health care delivery and insurance functions. It then tries to enhance efficiency of service provision internally, via the incentives it creates for the employed doctors.

Crucially, two important attributes of health plans are not measured in the data: quality of care and cost-sharing requirements. Omission of these variables is not a specific failure of the HCFA data. Rather, these attributes are intrinsically difficult to measure. As noted by Blumenthal (1996), “Experts have struggled for decades to formulate a concise, meaningful, and generally applicable definition of the quality of health care.”

Similarly, the cost-sharing rules of Medicare and Medigap plans are quite complex, with co-pays and deductibles contingent on condition, treatment and provider. The Center for Medicare and Medicaid Services (2015) (CMS) guide “Medicare and You” ran to over 150 pages. Section 3 alone, entitled “Find out if Medicare covers your test, service or item” is 33 pages long. Yet it is far from complete; the 2nd page of Section 3 states “copayments, coinsurance, or deductibles may apply for each service listed on the following pages. Visit Medicare.gov or call 1-800-MEDICARE to get specific cost information.” Given this complexity, it is obviously difficult to construct an overall measure of the cost-sharing requirements of Medicare. Similar problems apply to other plans.⁸

The difficulty of constructing quality and/or cost-sharing measures is an important problem, as these may be key factors in insurance choice. However, a key aspect of the HCFA data is that it contains attitudinal data in which consumers are asked how important it is to them that an insurance plan possesses certain attributes. The questions and response frequencies are shown in Table 2.

Economists typically eschew attitudinal data, on the grounds that it tells us nothing about consumers’ (monetary) willingness to pay for product attributes. But HK showed that these data are strong predictors of choice behavior. Specifically, HK showed how attitudinal data can be combined with observed health plan choices to measure both: 1) how consumers value unobserved attributes, and 2) the *perceived* levels of unobserved attributes for each plan

⁸ In principle, one could take the rules of a health plan, integrate over the distribution of possible health events for an individual, and construct expected out-of-pocket health care costs. This is a complex calculation, and, as we see, it implies that cost sharing is actually person/plan specific.

in the market. The idea is to treat the responses to attitudinal questions as noisy indicators of consumer preferences when estimating a choice model.

The insurance choice model in HK is specified as follows: Let X_j denote the vector of observed attributes of insurance option j , where $j = 1, \dots, 5$ indexes the five health plan options listed in Table 1. The attributes in X_j are the listed in the five rows of Table 1.⁹ Next, let A_j denote the vector of unobserved attributes of insurance option j . In this case these are: (1) Cost Sharing and (2) Quality. Then, letting U_{ij} denote expected utility to person i if he/she chooses insurance option j , we have:

$$(1) \quad U_{ij} = X_j \beta_i + A_j W_i + \varepsilon_{ij}$$

Here β_i is the vector of utility weights that person i attaches to the observed attributes, while W_i is the vector of utility weights that person i attaches to the *unobserved* attributes. The stochastic term ε_{ij} is assumed *iid* type I extreme value, giving a multinomial logit model.

In conventional choice modelling we learn about the person-specific utility weights β_i and W_i solely by observing choice behavior. This is what we will refer to as a “pure revealed preference approach.” But the innovation in HK is to show that the attitudinal measures described in Table 2 can give us important *additional* information about β_i and W_i . Specifically, HK code the responses to the attribute importance questions as 1 for “doesn’t matter,” 2 for “like to have,” and 3 for “have to have.” Then, letting:

S_{ik} = the importance (1, 2 or 3) that person i says he/she assigns to attribute k ,

β_{ik} = the utility weight that person i truly attaches to observed attribute k ,

they assume that the utility weights on the observed attributes are given by:

⁹ Note that premiums are measured in \$ per month, while Drug coverage, Preventive care, Provider choice (a 0/1 indicator) and Submit Claims are 0/1 indicators.

$$(2) \quad \beta_{ik} = \beta_{0k} + \beta_{1k} S_{ik} + \mu_{ik}$$

where β_{0k} and β_{1k} map the S_{ik} into utility units, and μ_{ik} is “measurement error.” Thus, the HK model allows for the possibility that consumers who say they value an attribute highly also act as if they value it highly. If that is true, we should obtain $\beta_{1k} > 0$ if an attribute is “good,” and $\beta_{1k} < 0$ if the attribute is “bad.” On the other hand, if the attitudinal data is not useful for predicting behavior the slope parameters in (2) will be insignificant and close to zero.¹⁰

Finally, HK assume the utility weights on the *unobserved* attributes are given by:

$$(3) \quad W_{ip} = W_{1p} S_{ip}^* + v_{ip} \quad p=1 \text{ (cost share), } 2 \text{ (quality).}$$

This is analogous to (2), except that S_{ip}^* denotes person i 's stated importance for *unobserved* attribute p , while the slope coefficient that maps stated attribute importance into true attribute importance is now denoted W_{1p} , and the measurement error term is now denoted v_{ip} .¹¹

The intuition for how the HK model identifies the unobserved attribute levels A_j is straightforward. Consider an unobserved attribute like quality. Quite simply, HK infer that an alternative has high *perceived* quality if, *ceteris paribus*, people who say they care a lot about quality tend to pick that alternative. This implies $W_{12} > 0$ and $A_2 > 0$. Conversely, if the stated importance of quality is not predictive of behavior ($W_{12} = 0$) it is impossible to estimate the perceived quality levels of each alternative (so A_2 is not identified). As HK explain, this also means an intercept is not identified in (3).¹² Appendix A contains details of model estimation.

In preliminary analysis, HK tested the predictive power of the attitudinal data by

¹⁰ The HK model does not assume *a priori* that the attitudinal data is a good predictor of individual preferences. If the attitudinal data are uninformative, the slopes in (2) will be close to zero, and the intercept terms in (2) will tell us the average utility weights that consumers place on each attribute.

¹¹ HK assume that the measurement error terms μ_{ik} in (2) and v_{ip} in (3) have normal distributions. The variances of these distributions are additional parameters that must be estimated as part of the model.

¹² This is because even average utility weights on unobserved attributes are not identified if the attitudinal data is uninformative about preferences (see footnote 9).

estimating a simple multinomial logit with the five observed attributes in Table 1 as covariates. They then added interactions between the observed attributes and the stated attribute importance measures. The improvement in fit was dramatic, with the pseudo-R² roughly doubling. These simple results imply that attitudinal data (or psychometric data more generally) do provide useful information about preferences.

This finding is good news for the overall research program proposed in this chapter, which at its core involves: (i) testing whether consumers make “good” choices in complex environments and (ii) learning how we can help them make better choices. If one maintains a pure revealed preference approach to choice modeling, the question of whether a choice is “good” has no meaning.¹³ Thus, it is difficult to see how we can make much progress in this area unless we are willing to use attitudinal data (or psychometric data more generally) to help assess how well actual choices align with “true” preferences and attributes of products.¹⁴

Notably, marketers have been using various types of psychometric data to model and predict consumer demand for many years – see McFadden (1986), Louviere (1988), Hensher et al., (1999), Louviere et al., (2000), McFadden et al., (2002), Swait and Andrews (2003).

This work has passed a market test, in that it is widely used by actual firms to predict

¹³ This statement is tautologically true given that we define the “pure revealed preference approach” as that which takes observed choices as the outcome of rational decision making under full information. Of course, in some contexts revealed preference analysis can be used to test if observed choices are rationalizable – or whether they instead exhibit GARP violations (see Afriat, 1967). We discuss this in detail in Section 4.

¹⁴ Beshears et al., (2008) discuss some alternative strategies.

First, one might impose enough structure on the problem that one can infer “true preferences” from mistake laden observed choices. This means specifying a structural model that incorporates cognitive biases. In the context of insurance, one example of this approach is Abaluck and Gruber (2009), who we discuss in Section 4. In the context of labour supply, there is work by Fang and Silverman (2004) and Chan (2014), who estimate the degree of present bias of welfare program recipients. In general, however, we are sceptical of the scope for estimating “true preferences” from observed choices *alone* in the absence of direct measures of preferences and/or consumer information. For instance, Abaluck and Gruber (2009) require strong assumptions about the functional form of utility and formation of expectations, while Fang and Silverman (2004) and Chan (2014) require quite special variation in the data (i.e., imposition of welfare time limits).

Second, one might attempt to infer true preferences from choice contexts where cognitive biases are likely to be small. In our view this approach suffers from a number of difficult issues, including great reliance on the subjective judgement of the researcher about when such contexts arise, problems of external validity, and limited scope for application. As Beshears et al., (2008) note, “It would be strange to try to infer someone’s normative preferences without at least considering their own stated views on the question.”

demand. Ironically, a dogmatic adherence to the revealed preference paradigm raises the conundrum of why firms would squander so much money on psychometric market research.

Returning to the structural model, Table 3 presents estimates of equation (2), which describes how people value the observed attributes of insurance plan options. The estimates imply that the attitudinal data is highly predictive of individual level preferences. For each of the five observed health plan attributes, the slope coefficient mapping the attitudinal measures into true attribute importance weights is significant and has the expected sign.¹⁵

As an example, Table 4 details the model's prediction of the utility weight that a person puts on drug coverage, depending on how much the person says he/she cares about this attribute. Notice that the utility weight ranges from a low value of 0.441 if the person says the attribute "doesn't matter," to a high value of 1.209 if the person says it is an attribute that he/she would "have to have." Thus, consumers who say they "have to have" drug coverage act as if they place nearly 3 times as much value on that attribute as the consumers who say this attribute "doesn't matter." But does a coefficient estimate of 1.209 mean that these consumers care a lot about drug coverage?

In a choice model, the best way to interpret magnitudes of the coefficient estimates is to look at what they imply about how changes in plan attributes would affect market shares. We report such simulation exercises in Appendix B. To summarize, the model implies that senior citizens place very high values on provider choice and drug coverage, and only very modest weight on other observed attributes such as premiums.

Finally, Table 5 presents the estimates of equation (3), including the unobserved attribute levels (A_j) for each insurance plan. Consider first the estimates of quality of care. It is worth noting that we can only measure quality of each plan relative to a base alternative, as only quality *differences* affect choices in the model. In Table 5, the quality of Basic Medicare

¹⁵ The improvement in the log-likelihood function when the stated attribute importance measures are included in the model is over 100 points (from -1956 to -1834), a very dramatic improvement.

is normalized to zero, so it is the baseline. Thus, the positive estimates of A_2 for plans 2 and 3 imply that consumers perceive these plans as higher quality than Basic Medicare. This makes sense, as options 2 and 3 are Basic Medicare plus Medigap insurance that covers additional services. Thus, care under these options should be at least as good as under Medicare alone.

The negative estimate of A_2 for the IPA plan implies that consumers perceive the care provided under this plan as relatively low quality. In contrast, consumers perceive the care provided under the group HMO plan as better than under Basic Medicare (but not as good as under Basic Medicare plus either Medigap plan).

The results for cost-sharing requirements are rather surprising. As we see in Table 5, the estimates of A_{21} through A_{51} are all negative. Since the preference weight that multiplies this attribute is a preference for “low cost sharing,” a negative attribute level means that the plan requires more cost sharing than the base alternative (Basic Medicare). Thus, these estimates imply that the survey respondents perceive every alternative health insurance plan as having greater cost-sharing requirements than Basic Medicare. In fact, Basic Medicare has the highest cost-sharing requirements of any option.

At this point, it’s worth recalling the intuition for how we can identify the levels of the unobserved plan attributes. Basically, if people who say they care a lot about low cost-sharing tend (*ceteris paribus*) to choose a particular plan, it implies the plan is perceived as having low cost-sharing. Thus, as the people who say they care most about low co-pays are also the most likely to choose Basic Medicare, the HK estimates imply that people perceive Basic Medicare as having relatively low co-sharing.

As we emphasized earlier, it is difficult to form an overall measure of cost-sharing requirements. Nevertheless, we know Basic Medicare has the highest co-pays of any plan. This is unambiguous, as plans 2 to 5 all cover “gaps” in Medicare coverage.¹⁶ Thus, it seems

¹⁶ Obviously, we can’t form an objective ranking of plans 2 to 5 on the cost-sharing dimension.

clear that respondents have fundamental misperceptions about cost-sharing.^{17, 18} This result illustrates how the HK framework allows us to test if consumer perceptions are accurate.

The HK results are consistent with a substantial body of work in health services research finding senior citizens have important mis-perceptions about Medicare in particular and the supplemental insurance market in general., A number of survey studies have asked people their beliefs about both Medicare and supplemental insurance plans. These studies consistently find that people have major misperceptions about health plan coverage and rules. See, for example, Cafferata (1984), McCall et al., (1986), Davidson et al., (1992), Blendon et al., (1998), Kaiser Family Foundation (2000), McFadden (2006), Kling et al., (2008), Abaluck and Gruber (2009), Maestas et al., (2009).

Publications that explain Medicare and Medigap rules are readily available, but many studies find that seniors have difficulty understanding these materials (see Gibbs et al., 1996; Feldman et al., 2000; Harris-Kojetin et al., 2001; McCormack et al., 2001; Kolstad and Chernew, 2007; Harris and Buntin, 2008).¹⁹ This difficulty is not surprising – as we noted earlier, the CMS guide “Medicare and You” is roughly 150 pages long. And the CMS guide “Choosing a Medigap Policy” was 100 pages long in 2006.²⁰ The complexity of Medicare and Medigap rules seems to preclude explaining them in a concise way, despite the best efforts of CMS.

¹⁷ An alternative hypothesis is that people with low incomes may place a great weight on low co-pays, but that they simply cannot afford Medigap. We find this story implausible for two reasons. First, HK dropped respondents who used Medicaid, the medical insurance program for the poor, or who had SSI disability benefits, or who couldn't pay the Medicare Part B premium of \$28 per month. Thus, the poorest respondents are not represented in the data. Second, the HMO options only cost a little more than Basic Medicare, so it seems implausible that liquidity constraints would preclude those options.

¹⁸ Interestingly, the HK estimates do not imply consumer misperceptions about the five observed plan attributes. That is, consumers who say they care a lot about premiums, provider choice, etc. do act as if they place a relatively high weight on those attributes. Why would mis-perceptions be more important for cost-sharing requirements? Our hypothesis is that cost-sharing is simply much harder to understand. In contrast, plan attributes like provider choice are more evident “up front” (e.g., both the premium and whether one has to choose a doctor from a list are evident when one joins a plan).

¹⁹ There is a parallel literature showing that younger workers also have difficulty choosing among employer provided health plans (see, e.g., Chernew and Scanlon, 1998; Abraham et al., 2006).

²⁰ In 2007 this publication was compressed to roughly 50 pages, where it has remained since, but the evidence suggests it is still hard to digest.

Given the complexity of Medicare and Medigap rules, it seems likely that many senior citizens – particularly those with cognitive limitations – may have great difficulty making health plan choices. As a result, informational interventions aimed at helping them make better choices may be called for. Unfortunately, the literature has not reached clear conclusions on how to present health plan information so it is more easily understandable (see Spranca et al., 2000; Harris-Kojetin et al., 2001; Hibbard et al., 2002; and Uhrig et al., 2006 for steps in this direction). In particular, most studies find that plan choices are little affected by informational interventions.

Aside from this work on Medicare and Medigap, there is also a literature showing that younger consumers (i.e., under 65) also have difficulty understanding their health insurance plan options. See, e.g., Gibbs et al., (1996), Isaacs (1996), Tumlinson et al., (1997), Cunningham et al., (2001), Frank (2004), Bhargava et al., (2016). A recent paper in this literature is Handel and Kolstad (2015). They use data from a large employer where workers had a choice between two options: (i) a no-deductible “network HMO” or “preferred provider organization” (PPO) or (ii) a high-deductible catastrophic coverage plan (HD). For each plan, they construct the distribution of OOP costs for each worker using a sophisticated spending model. As in Harris and Keane (1998), they also obtain survey data that measures employee’s perceptions of the attributes of the health plans. This data reveals substantial misperceptions about the attributes of plans. For example, only 28% of HD enrollees and 16% of PPO enrollees know the maximum OOP cost under the HD plan. Only about 1/3 of PPO employees understood that the HD plan gave access to the same provider network.

In a pure rational choice framework, Handel and Kolstad (2015) argue that the choice between the HD and PPO plans would only depend on the distribution of OOP under each

plan, risk version and the plan premiums.²¹ Conditional on risk aversion, relatively healthy people (with low OOP risk) should choose the HD plan. But, as Handel and Kolstad find, to rationalize the data requires assuming a rather remarkably high level of risk aversion, as only 11 to 17% of workers choose the HD plan (in 2011 and 2012, respectively).

Next, adopting an approach similar to Harris and Keane (1998), the authors include perceived attribute measures in the insurance choice model. These turn out to be extremely predictive of behavior. For example, “consumers who believe that the PPO plan has a larger network of medical providers value the [HD plan] by \$2,326 less than someone who correctly knows that these plans grant the same access...” This confirms the findings of Harris and Keane that: (i) consumers place substantial weight on (perceived) provider choice when choosing a health plan, and (ii) misperceptions about plan attributes have a major effect on choice behavior.

2.1.2. Evidence that Consumers Fail to Properly Judge Insurance Costs and Benefits

Returning to Medicare, another way to gauge whether senior citizens understand the Medigap market is to test whether those people who can most benefit from having Medigap insurance are also the most likely to buy it. This is essentially what is done by Fang, Keane and Silverman (2008) – henceforth FKS. Basically, FKS used a very rich set of health measures to construct expected medical costs for each person in their combined HRS/MCBS data. Surprisingly, they found that people with *lower* expected medical costs were *more* likely to buy supplemental insurance – a phenomenon known as “advantageous selection.”

A possible explanation for advantageous selection is that healthier people may also be more risk averse with respect to out-of-pocket medical costs. But FKS and Fang et al., (2010) find no evidence to support this hypothesis.

Instead, FKS find that seniors with higher cognitive ability have greater demand for

²¹ This is because the two plans are equivalent in terms of provider network. Of course they may still differ (or be perceived as differing) on other dimensions such as convenience and reliability.

Medigap insurance, conditional on health status and other characteristics. This finding is quite significant, because standard rational choice theory assumes that all consumers have the ability to make optimizing choices. The idea that two otherwise identical consumers²² would choose different levels of insurance coverage simply because they have different levels of cognitive ability is hard to explain in this paradigm. The most straightforward explanation of the FKS results is simply that seniors with higher cognitive functioning are more aware of the fact that Basic Medicare leaves a large fraction of health care costs uncovered, and so they are more aware of the value of having supplemental insurance.

The Medicare Modernization Act of 2003 introduced a drug coverage component into Basic Medicare. The new benefit, known as Medicare Part D, took effect in 2006.²³ Part D drug insurance plans are sold by private insurers who negotiate prices with drug companies. The government role is: (i) to provide premium subsidies for low income enrollees, and (ii) to pay most drug expenses above a “catastrophic limit,” which in 2006 was \$5100.

Thus, a new private insurance market was created, with an array of Part D plans with different premiums and cost-sharing requirements.²⁴ In 2009 there were an average of 50 drug plans to choose among per CMS region- see Neuman and Cubanski (2009). Given this large choice set, CMS recognized they had created a complex choice situation. So they attempted to assist seniors by creating a “Medicare Plan finder” website.²⁵

McFadden (2006) and Winter et al., (2006) show that, given their health status, many consumers probably would have had lower drug costs under a different prescription drug plan from the one they actually choose. Maestas et al., (2009) looked at the prices of Part D plans

²² In FKS “identical” means equal health status, equal expected health care costs, equal income, equal levels of risk aversion, identical socio-demographics, etc.

²³ What we call “Basic Medicare” is the original program, which consists of Part A that covers hospital costs, and Part B that covers outpatient costs. Part C created the capitated Medicare HMOs.

²⁴ For instance, according to Neuman and Cubanski (2009), “... in 2009, patients with Alzheimer’s disease who were taking Aricept could have paid as little as \$20 for a month’s supply in one prescription-drug plan or as much as \$88 in another.”

²⁵ See <https://www.medicare.gov/find-a-plan/questions/home.aspx>.

that offered the same benefits, and found substantial dispersion in prices – again implying that many consumers were unable to find the best plan. Other papers in this literature are Kling et al., (2008), Lucarelli (2009) and Abaluck and Gruber (2009). These studies provide indirect evidence against the rational choice model, but the results could also be rationalized by standard models of incomplete information and search, particularly if we account for unobserved attributes of drug plans like customer service (Ketcham et al., 2015a,b).

In countries like the U.K. and Australia, that have national health care systems, supplemental insurance is not particularly important as a way to avoid uncovered costs. Instead, private health insurance (PHI) gives one access to a “parallel” private care system. This, in turn, enables one to jump queues for various types of treatment that have long waiting lists. Some conditions have much longer waiting times for treatment than others, so, if one’s health status makes a long-wait condition more likely, one has a greater incentive to buy PHI.²⁶

With this in mind, Johar et al. (2011) investigated the demand for private health insurance in Australia. Analogous to FKS, they used a rich set of health status measures to construct, for each person in their data: (i) the probability of needing treatment and (ii) the expected waiting time conditional on needing treatment. Analogous to FKS, they found that expected waiting time was actually *negatively* correlated with demand for PHI.

2.1.3. Evidence on Choice Set Complexity, Age and “Confusion”

The literature on how choice set complexity affects the level of consumer confusion is very limited, especially if we seek studies that look at this issue by age. One of the very few papers that attempt to directly address this issue is Hanoch et al (2009). They look at how consumers’ ability to understand attributes of health plans declines as the number of available plans is increased, differentiating consumers by age and other demographics.

²⁶ For instance, in the UK the waiting time for physical therapy is currently enormous. If one has a history of back, neck or shoulder pain, one had better have private insurance, or else be prepared to pay out-of-pocket for physical therapy.

Specifically, Hanoch et al (2009) use an experimental approach to study the role of cognitive limitations in Medicare drug plan choices. They randomly assign subjects to treatments with choice sets of either 3, 10 or 20 Medicare drug plans. The participants received a table that contained (i) total annual OOP cost under each plan, (ii) annual deductible, (iii) cost-sharing requirements, (iv) number of participating pharmacies, (v) distance to closest participating pharmacy, and (v) whether drugs can be obtained by mail order. These informational materials are similar to what consumers would actually have available to choose amongst drug plans, except the real world choice task is simplified in several ways. Most obviously, consumers are told their OOP cost under each plan, rather than needing to calculate it. And complex aspects of Medicare Part D like the “donut hole” were abstracted from. The respondents were then asked four (rather simple) informational questions about the plans (e.g., which plan had the lowest cost, the closest pharmacy, etc.). Respondents were paid \$10 per hour (to avoid any incentive to answer quickly), and the median time to complete the questionnaire was 45 minutes.

Hanoch et al. (2009) found that the ability to correctly answer the informational questions declined sharply with both (i) the size of the choice set, and (ii) the age of the subject. Across all treatments (i.e., choice set sizes), 74% of subjects could answer at least 3 of 4 factual questions correctly. But in a logit model that controlled for education, race, mental and physical health, the odds ratios for a successful outcome were 0.17 and 0.10 for choice set sizes of 10 and 20, respectively. Thus, the odds of successfully answering 3 of 4 questions fell by a factor of 10 when the choice set increased from 3 to 20.

Half the subjects were over 65 years of age. The odds ratio for a successful outcome was 0.42 for subjects who were over 65. Thus, older subjects had much greater difficulty in understanding the choice options, even controlling for measures of cognitive functioning. Unfortunately, the sample size (190 subjects) was too small to reliably estimate interaction

effects between choice set size and age. This is obviously an area where much further research is called for.

2.1.4. Effect of Choice Set Simplification on Consumer Choice and Market Equilibrium

We have discussed several papers that show how consumers seem to exhibit “confusion” when buying insurance, in the sense that the people who can most benefit from buying insurance do not appear to do so. Thus, confusion can dampen adverse selection, or even lead to the phenomenon of “advantageous” selection (see FKS), where relatively healthy people actually buy more insurance. In such contexts, Handel (2013) raises the question of whether interventions aimed at helping people make better choices might have the unintended consequence of exacerbating the problem of adverse selection in equilibrium. That is, if people who need more health insurance do tend to buy more comprehensive plans, the premiums of those plans will have to rise in order for insurers to continue to break even.

To address this issue Handel (2013) uses several years of data on a private firm that offered a set of several HMO and PPO options, and exploits a change in the menu of choices offered by the employer in the middle of the sample period to help identify switching costs. He estimates a choice model where consumers care about the mean and variance of OOP under each plan, and there is consumer inertia in switching. The plans are assumed identical on non-financial aspects (no horizontal differentiation). To obtain equilibrium, Handel posits a simple supply side model where plan premiums cover their costs plus a fixed markup, He reports simulations showing that, on net, reforms that make it easier for consumers to choose the best plan actually reduce welfare because the utility gain from better matching is outweighed by the utility loss from higher insurance prices.

This is an important finding, but we would argue that Handel (2013)’s analysis abstracts from one important consideration: if consumers are confused it lowers the price elasticity for all products by creating “artificial product differentiation.” This is an important

factor that would work toward lowering prices as consumers make better informed choices. It is worth noting that this “artificial differentiation” mechanism is operative in all markets for differentiated products where producers have some degree of market power, not just insurance markets. The assumptions of no horizontal differentiation and a fixed markup abstract from these issues.

Subsequent work has shown that the impact of confusion (also called “frictions” or “inertia”) on equilibrium prices is ambiguous, and that it depends on a number of features of the choice environment. Ericson (2014) presents evidence that consumer inertia raises prices in equilibrium in the Medicare Part D drug insurance market. Basically, he finds that plans charge low prices when they first enter the market, and gradually increase prices as they develop a client base of “inert” enrollees who are not very sensitive to price.

Spinnewijn (2016) shows that a reduction in frictions or consumers causes a flattening of the demand curve facing an individual insurance plan.²⁷ He shows that the effect of reducing confusion depends on the covariances among consumers’ risk type (or expected cost to the insurer), degree of risk aversion (or preferences more generally) and perception errors. Similarly, Handel et al., (2015) look at a stylized market with one plan. In numerical experiments they show that confusion reducing policies are most likely to be welfare enhancing if the mean and the variance of risk aversion is high, while the variance of costs (*ex ante* risk) is relatively low. Using the estimates from Handel and Kolstad (2015), which imply low risk aversion, their simple calibrated model implies a \$47 welfare loss per person, which is 99% of consumer surplus.

Polyakova (2015) does a similar sort of analysis using data from the Medicare Part D market. But, in contrast to the Handel (2013) and Handel et al., (2015) papers, which use a

²⁷ This follows from a simple selection argument. Suppose perceived value (or WTP) depends on expected cost plus a risk aversion term plus a friction (or perception error) term ($WTP = c + r + f$). Then if P is the market price we have that $E(f | WTP > P) > 0$ and $\partial E(f | WTP > P) / \partial P > 0$. Intuitively, consumers who are willing to pay very high prices for an insurance plan will tend to be (on average) making mistakes such that they are overvaluing the plan. Eliminating these mistakes would therefore flatten the demand curve.

simple “cost plus fixed administrative cost” pricing rule, Polyakova (2015) estimates a reduced form pricing rule based on observed pricing behaviour of insurers. She then estimates a structural discrete choice model of consumer demand. Finally, she combines the demand model and the reduced form pricing rule to solve for market equilibrium prices and quantities. This “quasi-structural” approach is very similar to that used by Ching (2010a, b) to study equilibrium in the market name brand and generic drugs.²⁸ Another key difference is that Polyakova’s choice model does not assume that plans differ only by ex-ante mean and variance of OOP. Rather she allows for consumers to have heterogeneous preferences over a number of horizontal attributes of drug plans. Finally, she distinguishes between preference heterogeneity and inertia, a distinction which our previous discussion suggests may be very important for welfare calculations.

Polyakova’s (2015) results are quite striking in that she finds very large consumer welfare gains from reducing confusion in the Medicare Part D market. When she shuts down inertia she estimates a \$455 per person increase in welfare due to better matching, which is 23% of annual drug spending. There is a further \$10 increase in welfare due to a modest *drop* in premiums. We conjecture this occurs because, in a market with horizontal differentiation, removing inertia raises the price elasticity of demand and this outweighs any upward price pressure due to increased adverse selection.

Ericson and Starc (2013) examine a specific mechanism for reducing consumer confusion: product standardization. They look at the Massachusetts Health Insurance Exchange (HIX), a program started in 2006 to help match uninsured individuals with health

²⁸ Interestingly, the issues studied by Ching (2010a, b) are fundamentally identical to those studied in the several papers discussed here. He noted the puzzle that many consumers remain loyal to brand name drugs even after identical but much lower priced generics become available. Furthermore, the brand names raise their prices at this point. This appears to be explained by the fact that the loyal customers who stay with the brand names even after low-priced generics enter are very insensitive to price, so the brand name firm faces smaller market share but less elastic demand after generic entry. These loyal name brand drug customers are, in a different context, behaving like the inertia bound consumers in the Ericson (2014) and Handel (2013) studies.

insurance plans.²⁹ Initially firms had wide latitude with respect to product offerings, and in 2009 there were 6 insurers offering 25 plans, many of which were differentiated in rather subtle ways. The HIX design was changed substantially from 2009 to 2010 in an effort to make it easier for consumers to see the differences between plans. In 2010 each insurer was required to offer 6 plans with different levels of coverage, with financial characteristics identical across all insurers (within each level). A choice platform made this structure transparent. Strikingly, the mean OOP declined by \$259 per year post-standardization, as consumers tended to shift to plans with more comprehensive coverage. Furthermore, adjusted for generosity, monthly premiums were roughly \$12 higher in 2010 (suggesting the greater salience of financials led to a slight worsening of adverse selection).

Ericson and Starc (2013) go on to fit separate choice models for both 2009 and 2010. They find that the parameters of the choice model change with the change of choice environment. In particular, the financial aspects of plans, which were now more clearly distinguished, became much more salient. The authors take the parameters from the second “simpler” environment as the “true” representation of preferences for welfare calculations. They then conclude that product standardization greatly increased welfare.

While not necessarily doubting this conclusion, we disagree about the correct interpretation of the change in parameters after the change in the choice environment. In our view, the correct interpretation is that the parameters we see in both models (before and after) are reduced form parameters that are functions of preferences, the choice set and the information platform. This is precisely why they changed. There presumably exist deeper structural parameters of preferences that would not have changed. A recurring challenge for this literature is how to properly evaluate welfare when our estimated decision rules reflect not true preferences but rather reduced form parameters that also vary with the choice

²⁹ This matching also served to aggregate individuals so they could buy insurance at lower group rates. Ericson and Starc (2013) look specifically at the unsubsidized part of HIX (called “Commonwealth Choice”) that dealt with people about 300% over the poverty line.

context.³⁰

2.1.5. Summary

In summary, the findings reviewed here have important implications for the design of “competitive” health insurance markets. As Angell and Kassirer (1996) note: “According to the theory, if consumers are given full information about the quality of the health plans they are considering, they will opt for higher-quality plans, or at least when they trade off quality for lower costs, they will be able to do so knowingly. In a competitive system, consumers can then vote with their feet – that is, change plans if they believe that they can obtain better quality for the same price....” But, as Hall (2004) notes: “to choose rationally across insurers [consumers] must be well informed about ... the plans offered. ... [but] many consumers ... have not had substantial experience in obtaining health care until they face ... illness.” The evidence that consumers have important misperceptions about their health insurance and health care options undermines a key tenet of the standard “choice is good” argument.

2.2. Evidence of “Confusion” in Making Health Care Choices

We turn next to the issue of how senior citizens, and consumers more generally, make choices about health care services (as opposed to health insurance). For instance, people are often faced with the need to make choices among alternative providers (i.e., physicians, surgeons, hospitals), alternative treatment options (e.g., surgery vs. non-invasive treatment), different drugs (e.g., brand name vs. generic – see, e.g., Ching, 2010a, b), elective tests (e.g.,

³⁰ In a context where consumers are uncertain or confused about true attribute levels (say, because of the number and complexity of alternatives) the true “structural” model might specify that choice depends on *perceived* attributes, and the econometrician might then attempt to estimate utility weights on perceived attributes. Unfortunately the econometrician can’t see perceived attributes. If he/she simply uses the true attributes instead, it creates an errors-in-variables problem. [Note: This is exactly the sort of problem that Harris and Keane (1998) or Erdem and Keane (1996) try to handle by allowing for a distinction between true and perceived attributes in choice models, in different contexts]. Given such a mis-specified model, if perceived attributes change we would expect the coefficients on actual attributes to change as well (precisely because those coefficients are reduced form functions of both (i) actual utility weights and (ii) the mapping between true and perceived attributes). To give a concrete example, if true and perceived premiums are uncorrelated, we would expect a zero coefficient on true premiums. If we improve information so that true and perceived premiums are highly correlated, the estimated coefficient on true premiums would presumably increase – moving closer to the structural utility weight. In Ericson and Starc (2013) this may well explain why the weights on financial characteristics increase in the 2010 model.

cancer screening – see, e.g., Fiebig et al., 2010; Keane and Wasi 2013), vaccination, and so on. Harris and Buntin (2008) give an excellent review of the substantial literature on this topic, so here we just highlight some key points.

A key problem is that is the quality of a physician or hospital, or the effectiveness of a treatment, is very difficult to measure. For instance, doctors can be graded based on process measures (e.g., what fraction of patients are screened for high cholesterol?) and/or outcome measures (e.g., what fraction of patients have cholesterol in a desired range?). Since 2004 the National Health Service in the UK has based 25 to 30% of physician pay on such measures (see Roland and Campbell, 2014). But the problems with such an approach to measuring quality are manifold. Which aspects of care or outcomes should be considered? And what perverse incentives are created? Will physicians be tempted to “teach to the test” and work to improve what is measured while neglecting other important aspects of health care quality?

Even if such problems can be overcome, and we develop measures of quality that make sense from an expert point of view, how can these measures be communicated to consumers in an understandable way? The understanding of quality measures requires a great deal of health related knowledge that few people possess. If a surgeon has a certain success rate in a certain type of operation, is that good or bad?

By analogy, the quality of a baseball batter can be well summarized by his batting average (BA), on-base percentage (OBP) and slugging average (SA). But if you are a person with only a passing knowledge of the game, and you are told a batter has BA=.251, OBP=.314 and SA=.386, you will have no idea what that means. In fact, only a person with a substantial knowledge of baseball and baseball statistics could interpret these figures.³¹

Given the difficulty of understanding health care quality measures, it is not surprising that most studies reviewed in Harris and Buntin (2008) find that people rely primarily on

³¹ In fact, these were the average values across all major league players in 2014. But no doubt more than a small handful of baseball aficionados would be aware of that.

factors like quality of personal interactions when choosing a doctor. It is not surprising that people tend to ignore technical information they do not understand, and instead rely on factors like inter-personal skills which they do understand. Harris and Buntin (2008) describe a number of experiments that attempt to present provider quality information to consumers in a more useful way, but success in this area has been limited.

3. Retirement Savings and Investment Planning

Next, we consider the evidence on whether people in general – and the elderly in particular – can understand the complex choices they face in regard to retirement planning. Standard economic models assume that people plan optimally for retirement. But if instead people have difficulty making decisions about retirement savings vehicles (e.g., pension, 401(k) or superannuation plans), we may see a growing population of senior citizens and elderly whose well-being is adversely affected by failure to plan ahead optimally.

Retirement planning influences macroeconomic income and productivity as well as individual welfare. As populations age, income and insurance provision for the elderly take a larger share of public funds, increasing the size of the public sector (Poterba 2014). The diminishing government investment, rising taxes and perverse labor market incentives that follow can reduce aggregate efficiency (James 1995). In addition, population aging can hamper entrepreneurship, making it less likely that rising productivity will compensate for slower growth (Liang et al., 2014). In that case it is important to develop policies to help people plan better for retirement right now.

3.1 Evidence of “Confusion” in Retirement Planning

In theory, efficient life-cycle planners should have hump-shaped lifetime wealth profiles, adequate retirement income, and judiciously chosen insurance against mortality, longevity and health shocks. In fact, there are striking inconsistencies between theoretical predictions and actual behaviour. Many households retire with inadequate savings, even

when contributions to plans are mandatory,³² the voluntary take up of longevity insurance is low (Mitchell et al., 2011), and many elderly decumulate at very modest rates (see, e.g., Guiso et al., 2002; Börsch-Supan, 2003; Milligan, 2005; Love et al., 2009; Poterba et al., 2011; Ooijen et al., 2014; Wu et al., 2015a). These outcomes are hard to reconcile with rational planning.

Strikingly, only 43% of surveyed American adults say they have ever tried to figure out what they need for retirement, including only 57% of 50 to 65 year olds (Lusardi and Mitchell 2011). Studies from across the developed world consistently find fewer than half of adults have attempted any financial planning for retirement (see, e.g., Alessie et al., 2011; Bucher-Koenen and Lusardi, 2011; Fornero and Monticone, 2011; and Agnew et al., 2013a).

To make good retirement savings decisions, consumers need both: (i) to know and understand the attributes of the products/services they are evaluating, and (ii) to possess the cognitive capacity and skills to make good choices among those products/services. They are likely to become confused if they don't have the facts about investment returns, survival, pension plan structures and government support to hand. They also need the basic numeracy, financial literacy, patience and personal efficacy to design and implement a plan. Empirical studies have highlighted both misperceptions about the key facts and serious deficits in the capacity of many people to make a plan and follow through. We will first discuss the question of whether people have adequate information and accurate perceptions, and then turn to the question of their cognitive capacity for planning:

3.1.1 Evidence that Consumers Hold Biased Expectations

There is clear evidence that many consumers hold biased expectations of variables that are critical to retirement planning, including investment returns, longevity and retirement dates. Subjective expectations of equity market returns show marked pessimism and

³² Skinner (2007) surveys evidence for and against inadequate retirement savings. Studies showing inadequate savings include, among others, Laibson et al., (1998), Mitchell and Moore (1998), Knoef et al., (2015), and for a more recent empirical scan see Poterba (2014).

heterogeneity, despite the fact that they are readily observed public information. For example, data from Dutch adults put the average expected one-year-ahead return to equities at 0.3% when the historical median rate of return was actually 14% (Hurd et al., 2011). Other studies show that returns expectations tend to track recent stock market performance, and severe crashes increase uncertainty and disagreement (Hudomiet et al., 2011). High subjective pessimism and uncertainty may explain low stock market participation by risk averse investors, which, in turn, could account for low lifetime investment earnings.

Similarly, many people are excessively pessimistic about their survival prospects. Numerous international studies find that people underestimate their life expectancy by around five years on average. These errors are larger for women and younger cohorts – groups who should anticipate living longer (Hurd, 2009; Wu, Stevens and Thorp 2015; Teppa and Lafourcade, 2013; Kutlu-Koc and Kalwij, 2013). Individuals also mis-estimate the shape of the survival curve, showing too much pessimism to near ages and too little at distant ages. This means that they are more likely to misjudge retirement consumption and longevity insurance decisions (Wu, Stevens and Thorp 2015; Teppa and Lafourcade, 2013).

In contrast, reported retirement intentions are optimistic compared with realized retirement outcomes. Hurd (2009) studied responses from the HRS showing that middle-aged people's subjective expectations of still working at age 62 were upward biased: the forecast rate of full-time work was 46% compared with a realized rate of only 32%. This difference between realizations and expectations persisted even up to within a year or two of the target age. People who expect to retire later and die sooner than they actually do are likely to save less than they would need to finance retirement consumption.

3.1.2 Evidence that Consumers Misunderstand Pension Plan Rules and Entitlements

The findings discussed in section 3.1.1 are perhaps not surprising, given the evidence already noted in Section 2 that that people have difficulties understanding probabilities in

general (e.g., Johnson et al., 1993; Peters, Hibbard et al., 2007). However, peoples' misunderstanding of retirement planning is not limited to probabilistic outcomes like returns or survival. It extends to objective quantities that can, in principle, be known with certainty:

For example, several studies have shown that many pre-retirees have a weak grasp of their pension plan rules and social security entitlements. Mitchell (1988) compared Survey of Consumer Finance responses of employees with administrative data on their pension plans, and found major gaps in what employees knew. For example, this included knowing whether their employer contributed to their DC account, as well as the rules governing early retirement. Similarly, Gustman and Steinmeier (2005) found that only about half of respondents could report an estimate of their pension and Social Security benefits, and that those who could often made large errors. Bottazzi et al., (2006) report a similarly large range of expectations errors by Italian workers around replacement rates.

Although superannuation is compulsory for almost all workers in Australia, mistakes about preservation ages – i.e., the age at which superannuation accounts can first be accessed – are common among middle-aged workers (see Agnew et al., 2013b). Similarly, less than one third understand the basic features of standard decumulation products like lifetime annuities (Bateman et al., 2015). The value of plan-specific knowledge rises with the stakes, and wealthier, older, higher income, better educated males and whites do tend to know more. But significant errors persist.

3.2 Evidence that Consumers Lack Financial Literacy

Hypothetically, suppose we could design informational interventions that would fill the gaps in knowledge that we have described. The question remains whether people would have the cognitive capacity and skills to engage in (near) optimal retirement planning. As is well-known, even simple versions of the theoretical life-cycle problem can only be solved using dynamic programming (DP) methods and substantial computing power - see Geweke

and Keane (2001). By contrast, fewer than half of adults in developed countries can correctly answer three questions about financial basics such as interest rates, inflation and risk diversification (Lusardi and Mitchell 2014). So, as with health insurance, the assumption that most people can make (near) optimal choices regarding objects as complex as pension plans and annuities does seem to strain credulity.³³

Cognitive ability and acquired human capital, in the form of financial literacy, are powerful influences on retirement welfare (Jappelli and Padula, 2013; Lusardi and Mitchell, 2014). For example, Dohmen et al., (2010) find that higher cognition is associated with more risk tolerance and patience, and hence more wealth. Banks et al., (2010) find that households with higher numeracy exhibit steeper rates of accumulation and decumulation of assets over the life-cycle, consistent with life-cycle theory. Poor numeracy and financial literacy are also related to low rates of stock market participation (Christelis et al., 2010; van Rooij et al., 2011), higher rates of mortgage delinquency and defaults (Gerardi et al., 2013), and higher rates of mistakes in processing investment risk (Bateman, Eckert, Geweke et al., 2016).

Unfortunately, however, measured numeracy among adults, like other forms of financial literacy, is generally weak. For example, in simple questions about proportions, percentages and probabilities, tests of Australian adults show median scores of two out of three correct answers (e.g., Bateman et al., 2015). Galesic and Garcia-Retamero (2010) report similar results for the US and Germany, finding that probabilities are particularly poorly understood. Consequently a large minority of people probably lack the skill to understand compounding and risk, concepts that are critical to savings and investment decisions.

Cognition varies within individuals over time as well as in the cross-section. Agarwal et al., (2009) find an inverse u-shaped pattern of financial skill that peaks in middle age. The

³³ As discussed in Geweke and Keane (2001) and Houser, Keane and McCabe (2004), optimal solutions of life-cycle problems can often be well approximated by simple (but clever) rules of thumb. So the issue is not really whether people can solve DP problems, but whether they can behave in a sophisticated enough way so as to approximate such a solution.

decline in cognition at older ages makes managing retirement increasingly hard for the very elderly. Stock picking and diversification skills of investors in their '60s and '70s drop off sharply compared with middle age (Korniotis and Kumar 2011), and rates of credit card mistakes rise (Agarwal et al., 2009). Perhaps even more concerning is the evidence that worsening cognition does not bring with it any less confidence in one's ability to manage finances (Gamble et al., 2014a). This makes the elderly especially susceptible to scams and fraud (Gamble et al., 2014b; Blanton et al., 2012).

Beyond general cognitive ability and numeracy, people need some specific skills to make and execute good savings plans. For example, an understanding of compounding is fundamental but not easy: only 18% of early baby boomers surveyed in the HRS could answer a simple question about compound interest correctly, with 43% of those who got it wrong giving a simple linear interest answer (Lusardi and Mitchell 2007). Administrative data, as well as laboratory and field experiments confirm individuals' tendency to linearize interest growth and so underestimate the benefits of long-term savings (Song et al., 2015; Stango and Zinman 2009).

Not only are interest calculations difficult for many people, but there is also evidence that many have problems even thinking about delayed payoffs. This is especially true for people who are prone to procrastination or who have a poor connection with their future self (Weber 2003; Ersner-Herschfield et al., 2009; Bartels and Urminsky 2011). As a result, people will delay, refuse or over-simplify long-term savings and investment decisions, like joining a pension plan, until some event triggers it, such as changing jobs. Others may be paralyzed by worry about making mistakes and incurring financial losses (Rangel 2005). Conversely, cognitive biases such as over-optimism or over-confidence can also lead to inaction, by creating an attitude that one is invulnerable and the future will take care of itself.

3.2.1 Evidence that Consumers Make Passive Choices

An important practical way to deal with procrastination and lack of financial planning ability is the use of automatic enrolment in retirement plans. Another is default settings for contributions and investment strategies. Defaults have been shown to have large and long lasting effects, especially on unsophisticated savers (Madrian and Shea, 2001; Beshears et al., 2009; Choi et al., 2002, 2003). They simplify a complex decision by reducing it to a comparison between the default and everything else, rather than a comparison between many possibilities. Defaults are sometimes also interpreted as an endorsement by an expert (Beshears et al., 2009). When asked why they choose defaults, many retirement plan members cite their own lack of skill for making a choice or their wish to delay a complicated task (Butt et al., 2015; Brown et al., 2015). In general, passive behavior channels operating through defaults are far more effective for increasing savings than incentives such as tax rebates that require active decisions (Chetty et al., 2012).

Not everyone procrastinates or lacks financial capability, but the fact that “nudges” such as default options are so effective is implicit evidence that many households avoid thinking about their future needs (or find the problem very hard). If making savings decisions can be a challenge, investment choices are even harder. The advanced normative theory of optimal portfolio allocations proposes highly individualized strategies consisting of complex dynamic hedges (see Bodie et al. (2009) for a survey). It goes without saying that unsophisticated investors can’t design and implement these investment programs on their own, and that default investment options will be, at best, rough approximations to the ideal. Even the simplest version of modern portfolio theory predicts that investors should choose a well-diversified portfolio to maximise expected risk-adjusted returns. However, each of these three factors, returns, risk and diversification, present challenges to naïve investors.

In regard to returns for example, there is evidence that retirement savers: fail to take

up matching offers that offer risk-free returns (e.g., Saez, 2009; Choi et al., 2011); fail to minimize fees that reduce expected returns when choosing between otherwise identical index funds (Choi et al., 2010); and make different decisions about investments depending on whether fees are shown as gains or losses (Hastings et al., 2010), whether returns are shown as long or short-term (Benartzi and Thaler 1999), and whether equivalent returns are shown as dollars and cents, ratios or percentages (Rubaltelli et al., 2005).

3.2.2 Evidence that Consumers are “Confused” by Investment Decisions

As noted above, many people cannot answer questions about basic probabilities correctly, so it is not surprising that long horizon investment risk is also hard to grasp. For example, most individuals cannot infer outcomes of repeated gambles, overestimating the probability of a loss. As a result, they make much higher allocations to stocks when shown the distribution of 30-year returns than that of 1-year returns (Benartzi and Thaler, 1999; Klos et al., 2005). Comparing changes in investment risk is difficult for many people. Bateman, Eckert, Geweke et al., (2016) observed modifications to retirement savings portfolio allocations of individual investors in an experiment where investment risk increased but returns stayed constant. They recorded that about 30% of allocation decisions violated basic expected utility axioms, indicating misunderstandings of increasing risk.

In general, investment decisions are susceptible to the way that risk is framed. So much so that Bateman et al., (2013), using portfolio allocation experiments, show that changing the way that investment risk for retirement accounts was described caused much more variation in allocation decisions than even a doubling of the actual volatility of investment returns.

More fundamentally, it is not clear exactly how “investment risk” is understood by retirement savers, but the conventional measure of volatility is probably not what most people have in mind. Portfolio theory emphasizes both upside and downside risk, but unsophisticated

investors may be more focused on losses. Some studies show that such perceived risk is a better predictor of asset choices than return variance (Weber et al., 2005).³⁴ Weber (2003) further argues that the abstraction and distance of the consequences of retirement savings decisions means that the affective (emotional) response needed to evoke action is often missing. Thus, retirement “risk” does not seem “risky.” Even setting aside the psychological distance between retirement investment decisions and their consequences, ordinary investors struggle to understand both what investment risk is and how it relates to returns.

Other studies show that unsophisticated investors know that diversification is a good principle but do not understand the risk-return trade-off. Many think that diversified portfolios actually have higher risk and higher expected returns than concentrated portfolios (see Weber et al., 2005 and Reinholtz et al., 2015). The widespread use of diversification heuristics further highlights misunderstandings. When investors are confronted with large, complex, investment menus, choices can degenerate into ad-hoc strategies. For example, people divide their wealth evenly between some or all investment options even though this actually reduces diversification (Benartzi and Thaler, 2001; Huberman and Jiang, 2006; Brown et al., 2007; Morrin et al., 2012; Agnew et al., 2011; Bateman, Eckert, Iskhakov et al., 2016; Bateman, Dobrescu et al., 2016). Overall, empirical studies of investment decisions by ordinary consumers show that people want higher returns and diversification but are confused about how to achieve them.

3.3 Can Disclosures, Education or Advice Reduce Retirement Planning “Confusion”?

Informational asymmetries, search costs, complexity of contracts and a lack of trust imply that financial markets are prone to failure (Campbell et al., 2011). Consumers’ lack of information, their cognitive limitations or their behavioral biases can exacerbate the effects of

³⁴ Other work has found that probability-weighted ranges of outcomes are better understood and result in fewer mistakes than information about negative return frequencies (Goldstein et al., 2008; Vlaev et al., 2009; Bateman et al., 2015). Fewer mistakes are unsurprising given the additional information ranges offer over negative return frequencies. Even so, regulators often stipulate that risk is reported as a likelihood of losses. See Bateman, Eckert, Geweke et al., (2016) for a review of regulator and industry use of risk framings.

market failure. This can mean that some households are more affected than others. Planning and investment mistakes are more common among poorer, less educated households; these households are also less likely to participate in risky asset markets (Campbell 2006). However the effects of mistakes or failures are not limited to one group of households. When unsophisticated households confront complex products with shrouded attributes, such as bank account or credit card fees, the outcome can be cross-subsidization from naïve to sophisticated households (Gabaix and Laibson, 2006). This can also limit financial innovation. Sophisticated households that tend to be early adopters of new products are also unlikely to forego the cross-subsidies from less aware consumers that extant products offer.

What is the solution? Is poor retirement planning a problem that could be solved by improved disclosure, simplified products, education or advice? Regulators wanting to minimize restrictions on consumer financial choices while ensuring some degree of protection have tended to rely on disclosure rules. However, this has often resulted in lengthy and complicated disclosures mainly designed to minimize legal risk rather than to communicate clearly.

More recently, regulators have drawn up templates for financial product disclosures that aim to make information easier for unsophisticated investors to understand.³⁵ But simplified disclosures do not ensure better decisions. For instance, tests of the way consumers use the Securities and Exchange Commission (SEC) summary prospectus find no significant difference over the long form of disclosure that it was designed to replace – except that consumers spent less time on their decisions. Investors in the experiment still did not

³⁵ The G20 endorsed a set of common principles on consumer financial services including the principle that consumers should be given information on the fundamental benefits, risks and terms of products. The U.S. Securities and Exchange Commission (SEC) adopted a new simplified (or enhanced) disclosure document for mutual funds in 2007 and the European Commission implemented a Key Investor Information document for collective investment vehicles, such as mutual funds, that aimed to make information comparable across jurisdictions and easy for consumers to understand. In Australia, the regulator, the Australian Securities and Investments Commission (ASIC) prescribed a short form for investment product disclosures limited to eight pages and fixed structures.

minimise fees and loads in their choices of mutual funds (Beshears et al., 2011). Similarly, tests of the short form disclosures in Australia (Bateman, Dobrescu et al., 2015) showed that retirement plan members focused on very few of the information items on the form, had difficulty integrating the information on returns, risk and asset allocation, and retreated to a simple diversification rule. In general, it is hard to predict how people will respond to simplifications and unintended consequences can follow (Navarro-Martinez et al., 2011; Agarwal et al., 2015).

As well as controlling information disclosure, regulators can attempt to simplify choices by standardizing products or setting defaults. These strategies tend to rely on passive choice rather than active use of disclosures. In Australia, for example, regulators concerned about complexity stipulated a standardized form for default retirement plans (called “superannuation funds”). They implemented “MySuper” regulations that aimed to support a simple, low-fee, scalable default structure. Default plan member contributions must be invested in a constant balanced or life-cycle (target date) portfolio (Super System Review 2010). Many plans have lowered fees and switched to life-cycle investment strategies in response to the standardization regime (Butt et al., 2016). Similarly, Keim and Mitchell (2016) showed that streamlining defined contribution pension fund investment menus reduced turnover rates, expense ratios and systematic risk factors of plan members, potentially improving savings outcomes.

Despite the observed correlation of knowledge and financial capability with improved wealth management and planning outcomes, it is not clear that general education is an effective solution to poor decision making either. Financial literacy is often measured by simple tests of objective knowledge about interest rates or inflation. However when confronted with novel and complex choices, people need “consumer expertise” (Fernandes et al., 2014). Expertise is the appropriate set of skills, knowledge, acquired experience and

psychological traits needed for specific financial decisions. General financial education, especially if it is not applied to an immediate decision, tends to decay quickly (Fernandes et al. 2014).

What about expert advice? Usage rates of financial advice vary across countries, but for the US, Sabelhaus et al., (2008) found 55 percent of retiring DC account-owning households use an advisor, either one they found on their own (42 percent) or one provided by their employer (13 percent). In theory, advisors can bring better knowledge and financial practices to unsophisticated clients, while giving them the benefits of economies of scale in skill and knowledge - see Hackethal et al., (2012). In practice however, advisors often have the role of both expert and salesperson, so conflicted incentives affect the benefit of their relationship with clients - see Inderst and Ottaviani (2009).

Empirical studies of advisor-client relationships show both costs and benefits. Broker directed mutual fund investments, for example, including investments in retirement portfolios, underperform direct investments or age-appropriate target date funds. They exhibit return chasing, higher risk and higher turnover (Bergstresser et al., 2009; Chalmers and Reuter 2012). Field studies of personal financial advice find evidence that advisors who gave poor quality advice were still trusted by clients, and that advisors confirm clients' biases to their own advantage (ASIC 2012; Mullianathan et al., 2012).

On the other hand, unbiased computer generated advice has been shown to result in better outcomes for few clients who take it up, typically wealthier and more sophisticated clients (Bhattacharya et al., 2012). Less educated and unskilled clients are less likely to pay for advice, but they are more likely to rely on it when they do (Hackethal et al., 2010; Holden, 2013). So, while advice holds promise for assisting retirement planning, the cheapest and most unbiased sources, such as online and automated advice, appear to be unattractive to the less sophisticated consumers who could potentially benefit the most.

In summary, it seems clear that many people make suboptimal preparations for retirement or do not plan at all. This is partly because they lack information, but also because the retirement planning problem is intrinsically difficult (e.g., it involves matters such as investment risk that are very difficult for many people to understand). As in the case of health insurance, retirement savings decisions are very complex, and consumers appear to exhibit confusion when making them.

4. Models of Choice Behavior that Incorporate Irrational Behavior and Confusion

As we have seen, there is a substantial body of work showing that consumers exhibit confusion when making choices about complex products like health insurance, health care and retirement planning. They both: (i) fail to understand key attributes of the choice objects, and (ii) exhibit choice behavior that appears to be non-optimal and subject to an array of behavioral biases. Examples include choice of dominated alternatives, sensitivity to default options, sensitivity of choices to the framing of information, delay, and so on.

As a result, traditional models of rational choice may not be adequate to understand or predict how consumers make choices in these areas. But, while there exists a large literature criticizing the rational choice paradigm, relatively few papers have developed positive (predictive) models of how people actually behave in very complex choice environments.

Thus, in this section, we discuss some attempts to build behavioral models of how people actually make choices about complex choice objects like health insurance and retirement plans. In general, these models depart from the standard “rational choice” paradigm and incorporate confusion and cognitive limitations into the choice process.

4.1. Allowing Perceived Attributes to Differ from True Attributes

In Section 2 we argued that the model of Harris and Keane (1998) was a promising approach to relaxing the assumption that people know and understand the attributes of choice alternatives perfectly. Instead, their procedure allows one to estimate the “perceived”

attributes of choice alternatives as distinct from the true attributes. To do this requires that one have available not just data on consumer choices but also auxiliary data on how consumers value (or rate) the attributes.

Subsequently, Harris, Feldman and Schultz (2002) – henceforth HFS – provided additional evidence on the validity of the HK methodology. They analyzed insurance plan choices of employed workers who were under 65, and hence not yet eligible for Medicare. They used data from the Buyers Health Care Action Group (BHCAG), a coalition of two-dozen employers in Minneapolis/St. Paul that contracts directly with health care providers. The employees of BHCAG member companies have a choice among several alternative health plans. They were surveyed about their plan choices in 1998, and they were also asked a series of questions about how much they valued various plan attributes.

A key aspect of the HFS study is that they pretended they did not observe premiums, in order to ascertain if the HK methodology could successfully uncover premium differences across plans by using data on survey respondents' stated importance of premiums. They found that true premium differences were accurately uncovered by the HK methodology. And, like Harris and Keane, HFS found that the use of stated attribute importance data led to dramatic improvements in model fit. These results are encouraging for the HK method. Substantively, the HFS results showed (yet again) that consumers pay relatively little attention to various measures of provider quality.³⁶

4.2. Relaxing Theoretical Constraints on Choice Model Parameters

Abaluck and Gruber (2009) proposed a way to incorporate “irrational” behavior into a standard choice model. In an application to Medicare Part D, they argue that when fully rational consumers compare drug plans they should only care about the level and variability of out-of-pocket costs (net of premiums), not the details of how this is achieved. To test this,

³⁶ Parente, Feldman and Christianson (2004) also use the HK approach to study health plan choices of University of Minnesota employees. I will not describe this work in detail, but simply note that they again find that attitudinal data is very predictive of consumer choices.

they estimate a choice model of the form:

$$(4) \quad U_{ij} = P_j\alpha + E(opc)_{ij}\beta_1 + \sigma_{ij}^2\beta_2 + c_j\beta_3 + Q_j\beta_4 + \varepsilon_{ij} \quad j = 1, \dots, J$$

Here P_j is the premium of plan j , $E(opc)_{ij}$ is expected out-of-pocket costs for person i under plan j , σ_{ij}^2 is the variance of out-of-pocket costs, c_j is a vector of plan financial characteristics that affect out-of-pocket costs, and Q_j is a vector of plan quality measures. The stochastic term ε_{ij} is assumed *iid* type I extreme value, giving a multinomial logit model.³⁷ Normative theory predicts: (1) that $\alpha = \beta_1 < 0$ because consumers should be indifferent between plans with equal values of net expected out-of-pocket cost, $P_j + E(opc)_{ij}$, conditional on risk, (2) that $\beta_2 < 0$, provided that consumers are risk averse, and (3) that $\beta_3 = 0$, as consumers should be indifferent among plan financial characteristics once one conditions on $E(opc)_{ij}$ and σ_{ij}^2 . Of course, rational consumers may also care about various plan quality measures ($\beta_4 \geq 0$).

The Abaluck-Gruber estimates indicate that $|\alpha| \gg |\beta_1|$, implying excessive sensitivity to premiums, $\beta_2 < 0$ but insignificant, giving only weak evidence of risk aversion, and $\beta_3 \neq 0$, implying that people do care about the particular assortment of features (e.g., premiums vs. co-pays vs. deductibles) by which a health plan achieves a given expected level and variability of out-of-pocket costs. They take these results as evidence against rational behavior.³⁸

While the Abaluck-Gruber approach seems intuitively appealing, Ketcham, Kuminoff and Powers (2015a) present some criticisms of their work that are worth considering. First, they argue there may be important omitted variables in (4). In particular, rational consumers may care about the identity of the firm offering a plan – i.e., a plan’s “brand name” – because

³⁷ This is a first order Taylor approximation to a CARA utility function.

³⁸ Another possible explanation of the $|\alpha| \gg |\beta_1|$ and $\beta_3 \neq 0$ result is that the consumers are using the financial rules of the plans to form $E(opc)_{ij}$ and σ_{ij}^2 via a different method from the econometrician. It is quite difficult to rule this out, or to determine if the consumers’ approach is superior or inferior.

some firms are perceived as more reliable, less likely to dispute claims, etc..³⁹ The failure of the theoretical restrictions on coefficients may be due to such misspecification. For this reason they propose to add brand indicators to the vector Q_j in equation (4). But more general misspecifications of utility are also possible.

In order to examine this issue, KKP implement a revealed preference (RP) test which does not rely on a particular utility function. First, one must specify the set of plan attributes that consumers care about. Then, a person's behaviour cannot be rationalized if he/she chooses a dominated plan (i.e., one that is worse on *all* relevant attributes than another plan in his/her choice set).⁴⁰ As long as a person passes this (weak) RP test, there exists some utility function that can rationalize his/her behaviour.

To begin, assume that consumers only care about premiums, expected out-of-pocket costs and the variance of out-of-pocket costs. In that case, KKP find that 75% of consumers made dominated choices in 2006. This figure remains rather stable through 2010. However, when KKP assume that consumers also care about the brand name of a plan, they find that only 20% of consumers made dominated choices in 2006, and this fraction is again stable through 2010.⁴¹

A naïve response to these results would be to debate: (i) whether the 20% of irrational consumers detected by KKP is large or small, and (ii) whether allowing "brand" to be a relevant attribute of drug plans is "too generous" to the rationality hypothesis (in the sense that a fully rational consumer would perhaps be able to parse the true "quality" of a plan in a way that is more precise than just relying on brand name as a portmanteau signal).

³⁹ In principle, such plan differences should be captured by the quality measures Q_j already included in (4). But Abaluck-Gruber use the CMS "star" measures, which are thought to be weakly related to true quality – see Harris and Buntin (2008).

⁴⁰ Formally, plan A is dominated by plan B if A is *strictly* worse than B on at least one attribute, and *weakly* worse than B on all other attributes.

⁴¹ In an additional test, KKP consider a choice model that imposes the theoretical restrictions $\alpha = \beta_1, \beta_2 < 0, \beta_3 = 0$ on (4), while also including brand dummies. They find that this model forecasts out-of-sample at least as well as the Abaluck-Gruber model.

Apropos of both points, Bhargava et al (2016) look at a more controlled environment where a single insurer offers a large set of plans to employees of a private firm (thus eliminating brand as a confound) and where the plans only differ on four financial characteristics (thus also eliminating quality measures like network size from consideration). They nevertheless find that 55% of the employees made dominated choices. Interestingly, employees who were older, lower income, female or who had more health problems were more likely to choose dominated plans.

In our view the most important point of KKP and Bhargava et al., (2016) is they provide clear evidence that consumer behaviour is heterogeneous. Both studies find that a significant fraction of consumers (i.e., 20% or 55%) behave quite irrationally, in that they make dominated choices. And, presumably, there is another group of consumers who make choices that can be rationalized, but only using utility functions that exhibit attribute trade-offs most of us would consider “odd.” A well specified econometric model should account for such heterogeneity in behaviour.

The obvious problem with (4) is that it assumes homogeneous consumers. A naïve test of the theoretical restrictions $\alpha = \beta_1, \beta_2 < 0, \beta_3 = 0$, is in fact a test of a complex joint hypothesis: (i) coefficients are homogenous across consumers, (ii) the theoretical restrictions hold for all these homogeneous consumers, and (iii) as KKP note, there are no other types of misspecification (e.g., omitted variables). Notably, given heterogeneity in parameters, the theoretical restrictions that $\alpha = \beta_1, \beta_2 < 0, \beta_3 = 0$, could hold for every consumer in the sample, but be violated in the pooled data.⁴²

4.3. Allowing for Heterogeneity in the Choice Process

In our view, a promising approach to this type of problem is a model of “process heterogeneity.” This builds on and extends earlier work by El-Gamal and Grether (1995),

⁴² This is a version of a point made in Keane and Runkle (1990). Even if every person in a sample is making rational forecasts, the pooled data will generate evidence of irrational behavior.

Geweke and Keane (1999, 2000, 2001), Houser et al., (2004) and Keane and Wasi (2013).

For example, consider a model with two types of people, a rational type and a non-rational type:

$$(5a) \quad U_{ij} = \{E(opc)_{ij} + P_j\} \beta_{1i} + \sigma_{ij}^2 \beta_{2i} + Q_j \beta_{4i} + \varepsilon_{ij} \quad w.p. \quad p_1$$

$$(5b) \quad U_{ij} = P_j \alpha_i + E(opc)_{ij} \beta_{1i} + \sigma_{ij}^2 \beta_{2i} + c_j \beta_{3i} + Q_j \beta_{4i} + \varepsilon_{ij} \quad w.p. \quad 1 - p_1$$

Equation (5) says that a fraction p_1 of consumers are “rational,” and make decisions based on the utility function in (5a), while a fraction $1-p_1$ are “irrational” or “confused” and make decisions according to (5b). Equation (5a) incorporates the restrictions of rational choice theory as suggested by Abluck and Gruber, $\alpha_i = \beta_{1i}$, $\beta_{2i} < 0$, $\beta_{3i} = 0$, but at the *individual* level, while (5b) does not impose these restrictions.

Aside from allowing for two behavioral types, equation (5) also generalizes (4) by allowing for heterogeneity in utility function parameters within each type. We would not expect the parameter distributions to be the same for each types, so we might write:

$$(6a) \quad (\beta_{1i} \ \beta_{2i} \ \beta_{4i})' \sim N[(\beta_1^r \ \beta_2^r \ \beta_4^r)', \Sigma_1] \quad \text{if type} = 1$$

$$(6b) \quad (\alpha_i \ \beta_{1i} \ \beta_{2i} \ \beta_{3i} \ \beta_{4i})' \sim N[(\alpha^c \ \beta_1^c \ \beta_2^c \ \beta_3^c \ \beta_4^c)', \Sigma_2] \quad \text{if type} = 2$$

where the superscript “*r*” denotes rational while “*c*” denotes confused.

Finally, the stochastic term ε_{ij} is assumed *iid* type I extreme value in both (5a) and (5b). Thus, if we *condition* on a person’s type and his/her preference parameters, we have a simple multinomial logit model. But, given that we don’t observe a person’s true type and preference parameters, in order to form his/her likelihood contribution we must form the

unconditional probability of his/her choice by integrating over these unobservables.⁴³ This is closely analogous to how one forms the unconditional choice probabilities in the Harris and Keane (1998) model, which we describe in Appendix A, so we won't repeat the details here.

Estimation of the model (5)-(6) would give an estimate of the fraction of rational consumers in the population ($1-p_1$). It is important to note, however, that it would not categorize particular consumers as either rational or irrational. Rather, given the likelihood, we could construct the posterior odds, for each person in the data, that his/her behavior is described by (5a) or (5b). A useful specification check on the model in (5)-(6) is that we would expect consumers' posterior probabilities of being "irrational" to be closely related to whether they pass the rationality tests proposed by Ketcham et al., (2015a),⁴⁴ as well as to variables like cognitive ability that we would associate with decision making ability.

As we have discussed at several points in Sections 2-3, the question of how to do welfare analysis if consumers' "decision utility" departs from their "true utility" remains difficult and unresolved. One option that one might consider is to estimate a model of process heterogeneity as in (5)-(6) and then use the estimates for the "rational" type (i.e., the type whose estimated utility parameters obey theoretical restrictions) to perform welfare analysis. Such an approach relies on the theory restrictions being correct, and on the distribution of taste parameters among the sophisticated type being representative of the whole population.⁴⁵

⁴³ The model in (5)-(6) is a type of "mixed" logit model, with two stages of mixing. The individual level logit models are mixed using both (i) the mixing distribution determined by (6) and (ii) the type proportions p_1 and $1-p_1$.

⁴⁴ Ketcham et al., (2015b) propose to divide consumers into "rational" vs. "suspect" groups based on whether they pass the revealed preference test in Ketcham et al., (2015a), and whether they can answer a basic knowledge question about Medicare drug plans. They find the probability of being labelled "suspect" is systematically related to demographic variables that may proxy for cognitive ability (e.g. age, education, health status). They show that choice models like (4) estimated on the "rational" vs. "suspect" groups have very different parameters, with those for the "rational" group coming much closer to satisfying the restrictions suggested by Abluck and Gruber. They also show how to treat the parameters of the "rational" group as the true utility weights in evaluating welfare effects of policies aimed at helping the "suspect" group make better choices. The main limitation of their approach is they continue to assume homogeneous parameters within the "rational" and "suspect" types.

⁴⁵ In other words, the difference between the sophisticated and unsophisticated types lies in decision making ability, quality of information, and so on, but not in preferences themselves.

More generally, given estimates of the distribution of taste parameters as in (6), we could do many interesting exercises, such as: (i) assess the magnitude and welfare consequences of departures from rational choice behavior, (ii) assess the welfare implications of changing product attributes, or restricting choice sets, and (iii) predict the demand for (and welfare consequences of) introducing new products. Of course, the conclusions of such an analysis will depend on the parametric forms in (5)-(6). But it is not possible to do quantitative welfare analysis without parametric assumptions on utility.⁴⁶

4.4. Incorporating Confusion in the Choice Process

In the standard interpretation of the multinomial logit model due to McFadden (1974), the error terms ε_{ij} represent unobserved attributes of products for which consumers have heterogeneous tastes. For example, Blue Cross Blue Shield (BCBS) may have a high value of Q_j because it is widely perceived as high quality. But BCBS would only have a high ε_{ij} if person i has a personal reason for liking that brand (e.g., person i had a very good previous experience with BCBS). It is important to emphasize that consumer choice behavior is not “random” in the logit model. It only appears that way to an analyst who cannot observe ε_{ij} .

An important extension of our model is to allow for confusion in choice behavior. One way to capture confusion is to introduce genuine randomness into choice behavior. We propose to do this by adding a source of randomness to equation (5b). We then have:

$$(5b)' \quad U_{ij} = P_j \alpha_i + E(opc)_{ij} \beta_{1i} + \sigma_{ij}^2 \beta_{2i} + c_i \beta_{3i} + Q_i \beta_{4i} + \omega_{ij} \rho(A_i) + \varepsilon_{ij}$$

Here, $\omega_{ij} \sim N(0,1)$ captures a mistake in how consumer i evaluates the “true” utility that he/she will derive from choice of option j . The parameter $\rho(A_i) \geq 0$ is a scaling factor that

⁴⁶ Indeed, even “non-parametric” revealed preference analysis, which allows to test for rational behaviour (but not to do quantitative welfare analysis), is not fully “non-parametric” as it requires assumptions about what attributes of products do or do not generate consumer utility.

captures the magnitude of the consumer's mistakes.⁴⁷ A_i is a vector of both (i) individual characteristics, such as cognitive ability, financial knowledge, age, etc., that may influence a person's level of difficulty in making decisions, and (ii) contextual variables like size of the choice set or number of attributes, that influence the complexity of the choice situation.

The assumption that the scale of optimization errors is related to cognitive ability is motivated by the results of Fang et al., (2008). In their model of demand for health insurance they included a set of "behavioral" variables that would not be included in a typical rational choice model. Most importantly, they included a measure of cognitive ability. They found that, *ceteris paribus*, cognitive ability has a strong positive effect on demand for insurance. We hypothesize that people with higher cognitive ability are better able to understand the benefits of insurance, and better able to evaluate different plan options.^{48, 49}

The model in (5a), (5b)' and (6) can be used to study a number of different types of possible departures from rationality: First, the model generates an estimate of the proportion of rational consumers in the population (p_1). Second, by examining the estimates of $\rho(A_i)$ we can learn about the extent of "confusion" in choice behavior, as well as discovering whether some types of people exhibit more confusion than others.

Third, by looking at the *distribution* of the parameter vector $(\alpha_i \beta_{1i} \beta_{2i} \beta_{3i} \beta_{4i})$ we can learn a great deal about the nature of departures from rational behavior. For instance, do a large fraction of people have $|\alpha_i| \gg |\beta_{1i}|$, meaning they place excessive weight on premiums vs. out-of-pocket costs? Or are these excesses statistically significant but quantitatively quite

⁴⁷ The inclusion of the scale factor in the model is motivated by the work of Fiebig et al., (2010), who find strong evidence of scale heterogeneity in consumer choice behaviour.

⁴⁸ Fang et al., (2008) find not only that high cognitive ability people are more likely to buy insurance, but also that they tend to be healthier. Together, these two factors mean that healthier people are more likely to buy insurance. This phenomenon – which contradicts the standard prediction of adverse selection models – is known as "advantageous selection."

⁴⁹ A number of studies have also found that proxies for cognitive ability are related to the probability of rational decision making. Recent examples are Ketcham et al., (2015b), discussed earlier, and Bateman, Eckert, Iskhakov et al., (2016a, b), who study allocation of retirement funds to annuity vs. phased withdrawal plans, and find that measures of numeracy and financial literacy predict rational allocations.

modest? If the latter, this would cast the statistical finding that many people overweight premiums in quite a different light (as the problem would appear to be fairly modest).⁵⁰

Fourth, we can simulate the estimated model to learn how much choice behavior would be affected if the confusion term $\omega_{ij}\rho(A_i)$ were shut down. If we assume that the confusion term is not part of “true” utility, this exercise would allow us to assess the welfare loss due to confusion.⁵¹ Another interesting exercise is to simulate behaviour under a simpler menu of choice options than that which exists in the data. In a rational choice model restricting choice must reduce utility, but, in the presence of confusion, restriction (or simplification) of the choice set can potentially lead to an increase in consumer welfare.⁵²

Finally, note that other extensions of the process heterogeneity model are possible. For example, there is no reason to limit the mixture model to two types. One could add other processes like “always choose the default,” or other heuristics that are prevalent in the data.⁵³

4.5. Simple Methods to Model Dynamics that Relax Optimizing Assumptions

As we discussed in Section 3, optimal life-cycle planning requires the solution of a complex dynamic programming (DP) problem. But actual decision making in the domain of retirement planning often departs in obvious ways from normative principle, and people often seem to react to the difficulty of the problem with delays or procrastination. Next, we discuss models of choice behavior that accommodate such features.

⁵⁰ For example, we might find α^2 only slightly greater than β_1^2 but with the difference significant, and $Corr(\alpha_i, \beta_{1i})$ very high (but significantly less than one). Together, these results would imply statistically significant but economically negligible departures from rationality.

⁵¹ This is analogous to Kahneman et al., (1997)’s distinction between decision vs. hedonic utility.

⁵² To give a simple example, consider a choice set consisting of only two plans A and B. Plan B is dominated on all attributes observed by the econometrician. In the rational version of the logit model a tiny fraction of consumers do choose B just because they have very large values of ε_{iB} . However, if $\rho(A_i)$ is very large – indicating consumers are very confused by this choice – the probability of choosing B will be close to 50%. A paternalistic social planner could improve social welfare by banning plan B (provided of course that he/she does not put enormous weight on the small fraction of consumers with very large values of ε_{iB}).

⁵³ A precedent is Bateman, Eckert, Iskhakov et al., (2016b), who consider a model of process heterogeneity in allocation of funds to annuities vs. phase withdrawal plans. They let the mixture components include always choosing the default, making a 50/50 allocation, making a 100% allocation to one alternative (or the other), or drawing allocations from a beta-binomial., Their model differs fundamentally from (5)-(6), however, in that it is essentially descriptive (no option corresponds to optimizing behaviour). But they do find that lower cognitive ability people are more likely to use the naïve heuristics.

Interestingly, methods that appear to be relevant for retirement planning problems have already been developed in marketing research for the closely related problem of inventory planning. Specifically, Ching et al., (2009, 2014) – henceforth CEK – develop a model of consumer demand for a storable (or semi-durable) branded commodity. In this context, optimal behavior involves: (i) checking the prices of all brands of a product in every period, and (ii) solving a DP problem to determine both (ii-a) the reservation price for purchase of each brand, and (ii-b) the optimal quantity to buy in the event that the price of a brand is below its reservation price. Of course, the reservation prices and optimal purchase quantities both evolve in a complex way with inventories.

CEK argue that a normative model is unrealistic for two reasons: (i) For most products, consumers presumably do not have the time, interest or mental capacity to check all prices in every period, and (ii) in those periods when consumers do pay close attention to a category, they presumably make decisions using more or less sophisticated rules of thumb, not by literally solving a DP problem. Thus, CEK develop a two-stage model of demand for a storable branded product. In the first stage, consumers decide whether or not to pay attention to the product category. If they do decide to pay attention then, in stage two, they use a rule of thumb that may or may not provide a good approximation to the DP problem (depending on parameter estimates), as in Geweke and Keane (2000, 2001).

In CEK's empirical applications, the decision whether to consider a category is modelled as a simple probit or logit discrete choice model, where the factors that drive consideration are cues like advertising, displays and low inventory.⁵⁴ Their work was originally motivated by the observation that brand choice *conditional* on category purchase is very sensitive to price, while the decision to make a purchase in a category is quite insensitive to price. CEK showed that these seemingly contradictory facts could be explained if

⁵⁴ In the optimal solution consumers should consider a category in every period regardless of their inventory. Even if inventory is high, a low enough price would make it optimal to stock up even more.

consumers only occasionally look at (i.e., consider) a category.

It seems fairly clear how one might apply the CEK framework to financial products like annuities, life insurance or choice of retirement plans. As we discussed in Section 3, there is clear evidence that most consumers are averse to thinking about these products on a regular basis. For example, as is well-known, the typical consumer does not engage in a frequent re-balancing of his/her stock portfolio as the state of the world changes. It is natural to think of a framework where, in a first-stage, consumers decide on, say, a quarterly or annual basis whether to consider financial products in a certain category. The decision to consider could be driven by advertising, as well as by major life events such as retirement, children leaving home, a spouse passing away, selling a house and/or moving house, or reaching a milestone birthday. In the second stage, it would again be optimal to estimate a behavioral rule of thumb from the data, rather than imposing an optimal DP solution.

Given such an estimated model, one could simulate behavior under the model vs. under a normative solution to the planning problem. One could then evaluate whether or not the wealth losses from following the simplified decision process rather than the optimal DP solution are substantial (as in Houser et al., (2004)).

4.6. Summary and Directions for Future Work

To summarize, in this Section we have described three approaches to modelling choice behavior that depart in different ways from the standard rational choice paradigm. It is not difficult to see how these three approaches could be combined into a more general framework. Consider the demand for a private health insurance plan. One could adopt the CEK approach of assuming that consumers only occasionally consider such products, and adopt a first-stage consideration model that depends on factors like major life events. Then in the second stage, one could adopt a framework like our mixed logit model of equations (5)-(6), where some fraction of consumers make choices rationally while others do not, and there

is heterogeneity in the nature of departures from rationality.⁵⁵

Next, one could incorporate the Harris-Keane approach to relax the assumption that all product attributes are correctly observed by consumers. One could, for instance, treat expected out-of-pocket costs and product quality as latent variables in (5), provided that one had available measures of how much individuals value these attributes.

Finally, in dynamic choice contexts, where the value of each option includes not just a current payoff but also an expected stream of future payoffs (so that optimal choices would require a DP solution), we could, in the second stage, use a rule-of-thumb approximation to the value of each option, as in Geweke and Keane (1999, 2000) and Ching et al., (2014).

The approaches to modelling choice behavior that we have described in this section are “constructive” in the sense that they attempt to provide positive models of behavior that would be useful for prediction and policy analysis even if rational assumptions fail. We hope that such models will prove useful for empirical work in areas such as demand for financial products and insurance products, just as they have already proven useful in areas such as experimental economics, labor economics, marketing and health economics.

5. Conclusion

In this chapter we have reviewed evidence on how people make decisions in complex choice situations. This may refer to situations where the object under consideration is complex, in that it has many attributes, or some attributes that are difficult to understand or evaluate, and/or where the choice set itself is complex because there are a very large number of alternatives. We focus on three particular areas that are of special relevance to senior citizens: health insurance, health care, and retirement planning. The well-being of senior citizens depends critically on people making “good” choices in these areas, not just in old age

⁵⁵ Furthermore, as noted earlier, one could adopt additional behavioral modes, such as choice of defaults and other heuristics, as additional components of the mixture.

but over the whole life-cycle.

The idea that consumers are capable of making informed choices in markets for health care, health insurance, and retirement benefits rests on assumptions that they both (i) know and understand the attributes of the products/services they are evaluating, and (ii) possess the cognitive capacity and skills to make good choices among those products/services. Our review suggests that consumers in general, and the elderly in particular, fall far short of this ideal.

For instance, the evidence suggests that consumers have a very difficult time understanding the attributes of both public and private health insurance plans. Both the quality of care and expected out-of-pocket costs under alternative health care plans are very difficult for experts to measure, let alone for consumers to understand. There is ample evidence, for example, that most senior citizens in the U.S. do not understand cost sharing requirements of the Medicare plan, or how they are affected by supplemental insurance. The evidence also suggests that informational interventions have very modest impacts on health insurance and health care decisions.

A similar story holds in the area of retirement planning. Fewer than half of adults have attempted any financial planning for retirement. Consumers show misunderstandings and biases in critical areas of retirement planning knowledge, including life expectancy, investment returns and risk, retirement dates, pension ages and entitlements. Confusion extends to evaluating investment decisions and is exacerbated by low financial literacy. Improved disclosure, education and advice go some way towards improving outcomes, but impacts of education are often temporary, disclosure simplifications have unpredictable effects, and advice is affected by agency problems.

In light of these findings, we discuss ways to extend standard rational choice models to account for consumer confusion. For example, Harris and Keane (1998) develop an approach to choice modeling where perceived attributes may depart from true attributes. Geweke and Keane (2000, 2001) develop a method where consumers are assumed to use a

mixture of heuristics to solve a dynamic model. And Ching et al., (2009, 2014) develop a method that can account for consumer inattention, delay and procrastination.

In standard choice models in economics, the error term is assumed to arise from person specific tastes for the alternatives – tastes that the researcher can't see. Thus, choice behavior is deterministic from the point of view of the consumer. We have proposed a new type of choice model in which one component of the error term captures genuine randomness or confusion in consumer behavior. Under this “hybrid” view of the error term, randomness will arise primarily from taste heterogeneity in the case of simple products like laundry detergent. But as we move to more complex choices like health insurance or super funds, more randomness will be attributable to confusion.

Existing work has shown that randomness in choice due to confusion, and the extent of departures from rationality more generally, is moderated by non-traditional variables like cognitive ability, age, numeracy, education and financial literacy. These findings contrast sharply with traditional rational choice models where *all* consumers are assumed to be able to make optimal decisions, regardless of the complexity of the situation.

In all the models we have described, people make choices subject to optimization errors, so they do not base decisions on a normatively correct solution to the optimization problem. This leaves open the potential for carefully designed paternalistic interventions. We are however, quite cognizant of the potential for unintended consequences of such interventions.

Appendix A: Likelihood Function of the Harris-Keane Model

It is straightforward to estimate the Harris and Keane (1998) model in equations (1)-(3) using simulated maximum likelihood (SML). If the attribute importance weights β_i and W_i were known, the choice probability for a person would have a simple multinomial logit form. Since β_i and W_i are unobserved (we are estimating the parameters of their distribution), the simulated probability that person i chooses plan j is just the average over draws for β_i and W_i of multinomial logit choice probabilities:

$$(A1) \quad P(j | \theta, S_i, S_i^*) = D^{-1} \sum_{d=1}^D \left[\exp(X_j \beta_i^d + A_j W_i^d) / \sum_{k=1}^5 \exp(X_k \beta_i^d + A_k W_i^d) \right]$$

Here θ is the vector of all model parameters and S_i and S_i^* are the vectors of attitudinal measures for person i .

A technical point, explained at some length in Harris and Keane (1998), is that it is difficult to estimate both the scale of W_{1p} in equation (3) and the scale of the unobserved attribute levels A for each plan. To deal with this problem, Harris and Keane restricted W_{1p} to equal the inverse of the estimated standard deviation of the measurement error in equation (3), which, in turn, was restricted to be the same as the standard deviation of the measurement error in equation (2). Intuitively, these restrictions imply that the stated attribute importance measures are just as good at predicting peoples' preference weights on the unobserved attributes as they are at predicting peoples' preference weights on observed attributes.

Appendix B: Simulations of Harris-Keane Health Plan Choice Model

Given an estimated choice model, one can use it to simulate the impact of a change in plan attributes on the market shares of the various plans. One can also use the model to predict whether there would be substantial demand for new plans with particular attributes. Some examples of this type of exercise are provided in Table B.

Table B: Some Illustrative Experiments Using the Model

	Basic Medicare	Medigap w/o Drugs	Medigap w/ Drugs	IPA	HMO	“New Plan”
Baseline Market Shares	9.1%	9.4%	12.4%	25.6%	43.6%	
Medicare adds Drug Coverage	17.7%	8.2%	10.9%	22.2%	41.2%	
IPA adds Drug Coverage	6.7%	7.1%	9.1%	41.7%	35.5%	
IPA plan removes Provider Choice	11.4%	12.1%	16.3%	2.3%	57.7%	
Add “New Plan”	6.8%	7.4%	9.9%	19.6%	30.6%	25.8%

The first row of Table B reports a “baseline” simulation of the model – i.e., the model predictions of the market shares of the various plans. The second row of Table B reports the model prediction of what would happen if Basic Medicare were to add prescription drug coverage. The model predicts its market share would increase substantially, from 9.1% to 17.7%. Similarly, the 3rd rows shows that if the IPA were to add drug coverage its market share would increase from 25.6% to 41.7%. These results imply that many consumers find

prescription drug coverage to be a very attractive feature of a health plan.

The fourth row of Table B presents the model's prediction of what would happen if the IPA plan were to remove provider choice. The model predicts that its market share would dwindle to almost zero (2.3%). Thus, most consumers place great value on provider choice.

In other simulations (not reported), Harris and Keane (1998) found that moderate changes in premiums (i.e., \$20 per month increases) would have very small effects on plan enrollments. Thus, consumers appear to care greatly about provider choice and drug coverage, but not very much about premiums (at least within the modest range of premiums covered in the data).

In the bottom row of Table B, the model is used to predict what would happen if a new health insurance plan were introduced. The "New Plan" is designed to fill a gap that existed in the Minneapolis/St. Paul insurance market. Consider a segment of consumers who place a high value on provider choice and preventive care, but little value on prescription drug coverage. The plan best tailored to these tastes was the IPA. However, the IPA was perceived as being of very low quality (as well as having very high cost sharing), thus leaving these consumers without a very appealing option. The fact that so many people choose the IPA anyway (21.7%) suggests this configuration of preferences is rather common. The "New Plan" was designed to be like the IPA on observed attributes, but to have the same perceived quality as the group HMO ($A_{62}=.161$) and to have less perceived cost sharing ($A_{61}=-.150$).

The model predicts that the "New Plan" would be very popular, with a market share of 25.8%. Note that the "New Plan" differs from the group HMO primarily in that it allows provider choice but doesn't cover drugs. The model implies that a substantial segment of the population likes that option, provided it is also of reasonably high quality.

References

- Abaluck J., Gruber, J. (2009). "Choice inconsistencies among the elderly: evidence from plan choice in the Medicare Part D program". NBER Working Paper No. 14759 (forthcoming, *American Economic Review*).
- Abraham J., Feldman R., Carlin C., Christianson J., (2006). "Effect of Quality Information on Consumer Health Plan Switching". *Journal of Health Economics* 25, 762-781.
- Afriat, S.N. (1967). "The construction of a utility function from expenditure data." *International Economic Review* 8, 67-77.
- Agarwal, S., Chomsisengphet, S., Mahoney, N., Stroebel, J. (2015). "Regulating consumer financial products: evidence from credit cards". *Quarterly Journal of Economics* 130, 111-164.
- Agarwal, S., Driscoll, J. C., Gabaix, X., Laibson, D. (2009). "The age of reason: financial decisions over the life-cycle with implications for regulation". *Brookings Papers on Economic Activity*, Fall, 51-101.
- Agnew, J., Bateman, H., Thorp, S. (2013a). "Financial literacy and retirement planning in Australia". *Numeracy* 6, Article 7.
- Agnew, J., Bateman, H., Thorp, S. (2013b). "Superannuation knowledge and plan behaviour". *JASSA – The Finsia Journal of Applied Finance* 1, 45-50.
- Agnew, J.R., Szykman, L.R., Utkus, S.P., Young, J.A. (2011). "What people know about target-date funds: survey and focus group evidence". *Financial Security Project at Boston College Working Paper*, 2011-2.
- Alessie, R., Van Rooij, M., Lusardi, A. (2011). "Financial literacy and retirement preparation in the Netherlands". *Journal of Pension Economics and Finance* 10, 527-545.
- Angell, M., Kassirer, J. (1996), "Quality and the medical marketplace -following elephants". *New England Journal of Medicine* 335, 883-885.
- Australian Securities and Investment Commission (ASIC). (2012). "Shadow shopping study of retirement advice". Report 279, Sydney.
- Banks, J., O’Dea, C., Oldfield, Z. (2010). "Cognitive function, numeracy and retirement saving trajectories". *Economic Journal* 120, F381-F410.
- Bartels, D. M., Urminsky, O. (2011). "On intertemporal selfishness: how the perceived instability of identity underlies impatient consumption". *Journal of Consumer Research* 38, 182-198.
- Bateman, H., Dobrescu, I., Newell, B., Ortmann, A., Thorp, S. (2016). "As easy as pie: how retirement savers use prescribed investment disclosures". *Journal of Economic Behavior and Organization* 121, 60-76.
- Bateman, H., Eckert, C., Iskhakov, F., Louviere, J., Satchell, S., Thorp, S. (2013). "Financial competence, risk presentation and retirement portfolio preferences". *Journal of Pension Economics and Finance* 13, 27-61.

- Bateman, H., Eckert, C., Iskhakov, F., Louviere, J., Satchell, S., Thorp, S. (2015). "Individual capability and effort in retirement benefit choice". UNSW Australia Business School Research Paper No. 2014ACTL07, (forthcoming, *Journal of Risk and Insurance*).
- Bateman, H., Eckert, C., Geweke, J., Louviere, J., Satchell, S., Thorp, S. (2016). "Risk presentation and retirement portfolio choice". *Review of Finance* 20, 201-229.
- Bateman, H., Eckert, C., Iskhakov, F., Louviere, J., Satchell, S., Thorp, S. (2016). "Default and diversification heuristics in annuity choice". *Australian Journal of Management* published online 1 June 2016. doi:10.1177/0312896215617225.
- Bergstresser, D., Chalmers, J. M., Tufano, P. (2009). "Assessing the costs and benefits of brokers in the mutual fund industry". *Review of Financial Studies* 22, 4129-4156.
- Benartzi, S., Thaler, R. H. (1999). "Risk aversion or myopia? Choices in repeated gambles and retirement investments". *Management Science* 45, 364-381.
- Benartzi, S., Thaler, R.H. (2001). "Naive diversification strategies in defined contribution plans". *American Economic Review* 91, 79-98.
- Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C. (2008). "The importance of default options for retirement saving outcomes: evidence from the United States", *Journal of Public Economics* 92, 1787-94.
- Beshears, J., Choi, J.J., Laibson, D., Madrian, B.C. (2009). "The importance of default options for retirement saving outcomes: evidence from the United States". In: Brown J.R., Liebman J., Wise D.A. (Eds.), *Social Security Policy in a Changing Environment*, University of Chicago Press, Chicago.
- Beshears, J., Choi, J.J. Laibson, D., Madrian, B.C. (2011). "How does simplified disclosure affect individuals' mutual fund choices?" In: Wise, D.A. (Ed.), *Explorations in the Economics of Aging*. University of Chicago Press, Chicago.
- Bhargava, S., Lowenstein, G., Snyder, J. (2016). "Do individuals make sensible health insurance choices?" Forthcoming, *Quarterly Journal of Economics*.
- Bhattacharya, U., Hackethal, A., Kaelser, S., Loos, B., Meyer, S. (2012). "Is unbiased financial advice to retail investors sufficient? Answers from a large field study". *Review of Financial Studies* 25, 975-1032.
- Blanton, K. (2012). "The rise of financial fraud: scams never change but disguises do". Center for Retirement Research, Boston College Chestnut Hill.
- Blazer, D.G., Yaffe, K., Liverman, C.T. (Eds.) (2015). *Cognitive Aging: Progress in Understanding and Opportunities for Action*. National Academies Press.
- Blendon, R.J., Brodie, M., Benson, J.M., Altman, D.E., Levitt, L., Hoff, T., Hugick, L. (1998). "Understanding the managed care backlash". *Health Affairs* 17, 80-94.
- Blumenthal, D. (1996). "Quality of care - What is it?" *New England Journal of Medicine* 335, 891-894.
- Bodie, Z., Detemple, J., Rindisbacher, M. (2009). "Life cycle finance and the design of pension plans". *Annual Review of Financial Economics* 1, 249-86.

- Bonsang, E., Adam, S., Perelman, S. (2012). "Does retirement affect cognitive functioning?" *Journal of Health Economics* 31, 490–501.
- Börsch-Supan, A., (2003). *Saving: A Cross-National Perspective*. Academic Press, San Diego.
- Bottazzi, R., Jappelli, T., Padula, M. (2006). "Retirement expectations, pension reforms, and their impact on private wealth accumulation". *Journal of Public Economics* 90, 2187-2212.
- Brown, J. R., Farrell, A., M., Weisbenner, S. J. (2015). "Decision-making approaches and the propensity to default: evidence and implications". NBER Working Paper No.20949.
- Brown, J.R., Liang, N., Weisbenner, S. (2007). "Individual account investment options and portfolio choice: behavioral lessons from 401(k) plans". *Journal of Public Economics* 91, 1992 - 2013.
- Bucher-Koenen, T., Lusardi, A. (2011). "Financial literacy and retirement planning in Germany". *Journal of Pension Economics and Finance* 10, 565-584.
- Butt, A., Donald, S., Foster, D., Thorp, S., Warren, G. (2015). "Delegation, trust and defaulting in retirement plans: perspective from managers and members". CIFR Working paper 065/2015.
- Cafferata, G.L. (1984). "Knowledge of their health insurance coverage by the elderly". *Medical Care* 22, 835-847.
- Campbell, J.Y. (2006). "Household finance". *Journal of Finance* 61, 1553-1604.
- Campbell, J.Y., Jackson, H.E., Madrian, B.C., Tufano, P. (2011). "Consumer financial protection". *Journal of Economic Perspectives* 25, 91-113.
- Castle, E., Eisenberger, N.I., Seeman, T.E., Moons, W.G., Boggero, I.A., Grinblatt, M.S., Taylor, S.E. (2012). Neural and behavioral bases of age differences in perceptions of trust". *Proceedings of the National Academy of Sciences* 109, 20848-20852.
- Centers for Medicare and Medicaid Services (2006). "Choosing a Medigap policy: a guide to health insurance for people with Medicare". US Dept. of Health and Human Services.
- Centers for Medicare and Medicaid Services (2015). "Medicare and you". US Dept. of Health and Human Services.
- Chalmers, J., Reuter, J. (2012). "What is the impact of financial advisors on retirement portfolio choices and outcomes?" NBER Working Paper No.18158.
- Chan, M. (2014). "Welfare dependence and self-control: an empirical analysis", Working paper, University of Technology Sydney.
- Chernew, M., Scanlon, D.P. (1998). "Health plan report cards and insurance choice". *Inquiry* 35, 9-22.
- Chetty, R., Friedman, J.N., Leth-Petersen, S., Nielsen, T.H., Olsen, T. (2014). "Active vs passive decisions and crowd-out in retirement savings accounts: evidence from Denmark", *Quarterly Journal of Economics* 129, 1141-1219.
- Ching, A. (2010a). "A dynamic oligopoly structural model for the prescription drug market after patent expiration," *International Economic Review* 51, 1175-1207.

- Ching, A. (2010b). "Consumer learning and heterogeneity: dynamics of demand for prescription drugs after patent expiration," *International Journal of Industrial Organization* 28, 619-638
- Ching, A., Erdem, T., Keane, M. (2009). "The price consideration model of brand choice". *Journal of Applied Econometrics* 24, 393-420.
- Ching, A., Erdem, T., Keane, M. (2014). "A simple method to estimate the roles of learning, inventories and category consideration in consumer choice". *Journal of Choice Modelling* 13, 60-72.
- Choi, J. J., Laibson, D., Madrian, B.C., Metrick, A. (2002). "Defined contribution pensions: Plan rules, participant choices, and the path of least resistance". In: Poterba, J. (Ed.) *Tax Policy and the Economy* 16, MIT Press, Boston.
- Choi, J. J., Laibson, D., Madrian, B.C., Metrick, A. (2003). "Optimal defaults". *American Economic Review* 93, 180-185.
- Choi, J. J., Laibson, D., Madrian, B.C. (2011). "\$100 bills on the sidewalk: Suboptimal investment in 401 (k) plans". *Review of Economics and Statistics* 93, 748-763.
- Choi, J. J., Laibson, D., Madrian, B.C. (2010). "Why does the law of one price fail? An experiment on index mutual funds". *Review of Financial Studies* 23, 1405-1432.
- Christelis, D., Jappelli, T., Padula, M. (2010). "Cognitive abilities and portfolio choice". *European Economic Review* 54, 18-38.
- Cole, C., Laurent, G., Drolet, A., Ebert, J., Gutchess, A., Lambert-Pandraud, R., Mullet, E., Norton, M.I., Peters, E. (2008). "Decision making and brand choice by older consumers". *Marketing Letters* 19, 355-365.
- Cunningham, P.J., Denk, C., Sinclair, M. (2001). "Do consumers know how their health plan works?" *Health Affairs* 20, 159-166.
- Davidson, B. N., Sofaer, S., Gertler, P. (1992). "Consumer information and biased selection in the demand for coverage supplementing Medicare". *Social Science & Medicine* 34, 1023-1034.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U. (2010). "Are risk aversion and impatience related to cognitive ability?" *American Economic Review* 100, 1238-60.
- Eisenstein, E., Hoch, S. (2005). "Intuitive compounding: framing, temporal perspective, and expertise". Manuscript, Temple University.
- El-Gamal, M.A., Grether, D.M. (1995). "Are people Bayesian? Uncovering behavioral strategies". *Journal of the American Statistical Association* 90, 1137-1145.
- Enthoven, A.C. (1988). *The Theory and Practice of Managed Competition in Health Care Finance*. North Holland, Amsterdam.
- Erdem, T., Keane, M. (1996). "Decision-making under uncertainty: capturing dynamic brand choice processes in turbulent consumer goods markets". *Marketing Science* 15, 1-20.
- Erdem, T, Keane, M., Sun, B. (2008), "The impact of advertising on consumer price sensitivity in experience goods markets". *Quantitative Marketing and Economics* 6, 139-176.

- Ericson, K.M. (2014). "Consumer inertia and firm pricing in the Medicare Part D prescription drug insurance exchange," *American Economic Journal: Economic Policy* 6, 38-64.
- Ericson, K.M., Starc, A. (2013). "How product standardization affects choice: evidence from the Massachusetts Health Insurance Exchange," NBER Working Paper No. 19527.
- Ersner-Hershfield, H., Garton, M. T., Ballard, K., Samanez-Larkin, G. R., Knutson, B. (2009). "Don't stop thinking about tomorrow: individual differences in future self-continuity account for saving". *Judgment and Decision Making* 4, 280-286.
- Fang, H., Silverman, D. (2004). "On the compassion of time-limited welfare programs". *Journal of Public Economics* 88, 1445-70.
- Fang, H., Keane, M.P., Silverman, D. (2008). "Sources of advantageous selection: evidence from the Medigap insurance market". *Journal of Political Economy* 116, 303-349.
- Fang, H., Nicholas L., Silverman, D. (2010). "Cognitive ability and retiree health care expenditure". Michigan Retirement Research Center Working Paper 230.
- Fiebig, D., Keane, M., Louviere, J., Wasi, N. (2010). "The generalized multinomial logit model: accounting for scale and coefficient heterogeneity". *Marketing Science* 29, 393-421.
- Fernandes, D., Lynch Jr, J.G., Netemeyer, R.G. (2014). "Financial literacy, financial education, and downstream financial behaviors". *Management Science* 60, 1861-1883.
- Feldman, R., Christianson, J., Schultz, J. (2000). "Do consumers use information to choose a health-care provider system?" *Milbank Quarterly* 78, 47-78.
- Fornero, E., Monticone, C. (2011). "Financial literacy and pension plan participation in Italy". *Journal of Pension Economics and Finance* 10, 547-564.
- Frank, R. (2004). "Behavioral economics and health economics," NBER Working Paper No. 10881.
- Frank, R.G., Zeckhauser, R.J. (2009). "Health insurance exchanges—making the markets work". *New England Journal of Medicine* 361, 1135-1137.
- Gabaix, X., Laibson, D. (2006). "Shrouded attributes, consumer myopia, and information suppression in competitive markets". *Quarterly Journal of Economics* 121, 505-540.
- Galesic, M., Garcia-Retamero, R. (2010). "Statistical numeracy for health: a cross-cultural comparison with probabilistic national samples". *Archives of Internal Medicine* 170, 462-468.
- Gamble, K.J., Boyle, P.A., Yu, L., Bennett, D.A. (2014a). "Aging and financial decision making". *Management Science* 61, 2603-2610.
- Gamble, K. J., Boyle, P.A., Yu, L., Bennett, D.A. (2014b). "The causes and consequences of financial fraud among older Americans". Center for Retirement Research, Boston College, CRR WP 2014-13.
- Gerardi, K., Goette, L., Meier, S. (2013). "Numerical ability predicts mortgage default". *Proceedings of the National Academy of Sciences* 110, 11267-11271.
- Geweke, J., Keane, M. (1999). "Mixture of normals probit models". In: Hsiao, C., Lahiri, K., Lee, L.F., Pesaran, M.H., (Eds.), *Analysis of Panels and Limited Dependent Variable Models*, Cambridge University Press, Cambridge.

- Geweke, J., Keane, M. (2000). "Bayesian inference for dynamic discrete choice models without the need for dynamic programming". In: Mariano, R., Schuermann, T., Weeks, M.J. (Eds.), *Simulation Based Inference and Econometrics: Methods and Applications*, Cambridge University Press, Cambridge.
- Geweke, J., Keane, M. (2001). "Computationally intensive methods for integration in econometrics". In: Heckman, J.J., Leamer, E.E. (Eds.), *Handbook of Econometrics*, vol 5. Elsevier Science B.V.
- Gibbs, D.A., Sangl, J.A., Burrus, B. (1996). "Consumer perspectives on information needs for health plan choice". *Health Care Financing Review* 18, 55-73.
- Goldstein, D.G., Johnson, E.J., Sharpe, W.F. (2008). "Choosing outcomes versus choosing products: consumer-focused retirement investment advice". *Journal of Consumer Research* 35, 440-456.
- Grotz, C., Letenneur, L., Bonsang, E., Amieva, H., Meillon, C., Quertemont, E., Salmon, E., Adam, S. (2015). "Retirement age and the age of onset of Alzheimer's disease: results from the ICTUS study". *PloS one* 10, e0115056.
- Guiso, L., Haliassos, M., Jappelli, T. (2002). *Household Portfolios*. MIT Press, Cambridge.
- Gustman, A. L., Steinmeier, T. L. (2005). "Imperfect knowledge of social security and pensions". *Industrial Relations: A Journal of Economy and Society* 44, 373-397.
- Hackethal, A., Inderst, R., Meyer, S. (2010). "Trading on advice". CEPR Discussion Paper No. DP8091.
- Hackethal, A., Haliassos, M., Jappelli, T. (2012). "Financial advisers: a case of babysitters?" *Journal of Banking and Finance* 36, 509-524.
- Hall, J. (2004). "Can we design a market for competitive health insurance?" CHERE discussion paper No. 53.
- Handel, B. (2013). "Adverse selection and inertia in health insurance markets: when nudging hurts". *American Economic Review* 103, 2643-2682.
- Handel, B., Kolstad, J. (2015). "Health insurance for 'humans': information frictions, plan choice and consumer welfare." *American Economic Review* 105, 2449-2500.
- Handel, B., Kolstad, J., Spinnewijn, J. (2015). "Policy interventions in health insurance markets". NBER Working Paper No.21759.
- Hanoch, Y., Rice, T., Cummings, J., Wood, S. (2009). "How much choice is too much? The case of the Medicare prescription drug benefit". *Health Services Research* 44, 1157-1168.
- Harris, K., Buntin, M.B., (2008). "Choosing a health care provider: the role of quality information". *Research Synthesis Report* 14, 1-25, Robert Wood Johnson Foundation.
- Harris, K.M., Keane, M.P., (1998). "A model of health plan choice: inferring preferences and perceptions from a combination of revealed preference and attitudinal data". *Journal of Econometrics* 89, 131-157.
- Harris, K.M., Feldman, R., Schultz, J. (2002). "Measuring consumer perceptions of quality differences among competing health benefit plans". *Journal of Health Economics* 21, 1-18.

- Harris-Kojetin L.D., McCormack, L.A., Jael, E.F., Sangl, J.A., Garfinkel, S.A. (2001). "Creating more effective health plan quality reports for consumers: lessons from a synthesis of qualitative testing". *Health Services Research* 36, 447-476.
- Hastings, J., Mitchell, O. S., Chyn, E. (2011). "Fees, framing, and financial literacy in the choice of pension manager". In: Mitchell, O. S., Lusardi, A., (Eds.), *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*, Oxford University Press, Oxford.
- Hensher, D.A., Louviere, J., Swait, J. (1999). "Combining sources of preference data". *Journal of Econometrics* 89, 197-221.
- Hertzog, C., Kramer, A.F., Wilson, R.B., and Lindenberger, U. (2008). "Enrichment effects on adult cognitive developments. Can the functional capacity of older adults be preserved and enhanced?" *Psychological Science in the Public Interest* 9, 1-65.
- Hibbard, J.H., Slovic, P., Peters, E., Finucane, M.L. (2002). "Strategies for reporting health plan performance information to consumers: evidence from controlled studies". *Health Services Research* 37, 291-313.
- Holden, S. (2013). "When, why, and how do mutual fund investors use financial advisers?" In: Mitchell, O.S., Smetters, K., (Eds.), *The Market for Retirement Financial Advice*. Oxford University Press, Oxford.
- Houser, D., Keane, M., McCabe, K. (2004). "Behavior in a dynamic decision problem: an analysis of experimental evidence using a Bayesian type classification algorithm". *Econometrica* 72, 781-822.
- Huberman, G., Jiang, W. (2006). "Offering versus choice in 401 (k) plans: equity exposure and number of funds". *Journal of Finance* 61, 763 -801.
- Hudomiet, P., Kezdi, G., Willis, R.J. (2011). "Stock market crashes and expectations of American households". *Journal of Applied Econometrics* 26, 393-415.
- Hurd, M.D. (2009). "Subjective probabilities in household surveys". *Annual Review of Economics* 1, 543-564.
- Hurd, M., Van Rooij, M., Winter, J. (2011). "Stock market expectations of Dutch households". *Journal of Applied Econometrics* 26, 416-436.
- Inderst, R., Ottaviani, M. (2009). "Misselling through agents". *American Economic Review* 99, 883-908.
- Isaacs, S.L. (1996). "Consumer's information needs: results of national survey". *Health Affairs* 15, 31-41.
- James, E. (1995). "Averting the old age crisis: an international perspective". *Aging International*, June, 15-22.
- Jappelli, T., Padula, M. (2013). "Investment in financial literacy and savings decisions". *Journal of Banking and Finance* 37, 2779-2792.
- Johar, M., Jones, G., Keane, M., Savage, E., Stavrunova, O. (2011). "Waiting times for elective surgery and the decision to buy private health insurance". *Health Economics* 20:S1, 68-86.

- Johnson, E.J., Hershey, J., Meszaros, J., Kunreuther, H. (1993). "Framing probability distortions, and insurance decisions". *Journal of Risk and Uncertainty* 7, 35-51.
- Kaiser Family Foundation (2000). *National Survey on Americans as Health Care Consumers: An Update on the Role of Quality Information*, Menlo Park, California.
- Kahneman, D., Wakker, P., Sarin, R. (1997). "Back to Bentham? Explorations of experienced utility". *Quarterly Journal of Economics* 112, 375-405.
- Keane, M., Runkle, D. (1990). "Testing the rationality of price forecasts: new evidence from panel data". *American Economic Review* 80, 714-735.
- Keane, M., Wasi, N. (2012). "Estimation of discrete choice models with many alternatives using random subsets of the full choice set: with an application to demand for frozen pizza". Nuffield College Working paper 2012-W13, University of Oxford.
- Keane, M., Wasi, N. (2013). "Comparing alternative models of heterogeneity in consumer choice behaviour". *Journal of Applied Econometrics* 28, 1018-1045.
- Ketcham, J., Kuminoff, N., Powers, C. (2015a). "Which models can we trust to evaluate consumer decision making? Comment on choice inconsistencies among the elderly". NBER Working paper No. 21387, (forthcoming, *American Economic Review*).
- Ketcham, J., Kuminoff, N., Powers, C. (2015b). "Estimating the heterogeneous welfare effects of choice architecture: an application to the Medicare prescription drug insurance market". Working paper, Arizona State University.
- Keim, D. B., Mitchell, O. S. (2016). "Simplifying choices in defined contribution retirement plan design." NBER working paper No.21854.
- Kling, J.R., Mullainathan, S., Shafir, E., Vermeulen, L., Wrobel, M. (2008). "Misperception in choosing Medicare drug plans." Harvard University working paper.
- Klos, A., Weber, E. U., Weber, M. (2005). "Investment decisions and time horizon: risk perception and risk behavior in repeated gambles". *Management Science* 51, 1777-1790.
- Knoef, M., Been, J., Alessie, R., Caminada, K., Goudswaard, K., Kalwij, A. (2015). "Measuring retirement savings adequacy: developing a multi-pillar approach in the Netherlands". *Journal of Pension Economics and Finance* 15, 55-89.
- Kolstad, J., Chernew, M. (2009). "Quality and consumer decision making in the market for health insurance and health care services". *Medical Care Research and Review* 66, 28-52.
- Korniotis, G. M., Kumar, A. (2011). "Do older investors make better investment decisions?" *Review of Economics and Statistics* 93, 244-265.
- Kutlu-Koc, V., Kalwij, A. (2013). "Individuals' survival expectations and actual mortality". Discussion Paper No. 2013-013, Netspar.
- Laibson, D. I., Repetto, A., Tobacman, J., Hall, R.E., Gale, W.G., Akerlof, G.A. (1998). "Self-control and saving for retirement". *Brookings Papers on Economic Activity*, 91-196.
- Liang, J., Wang, H., Lazear, E. P. (2014). "Demographics and entrepreneurship". NBER working paper No. 20506.

- Louviere, J. (1988). "Conjoint analysis modelling of stated preferences: a review of theory, methods, recent developments and external validity". *Journal of Transport Economics and Policy* 22, 93-119.
- Louviere, J., Hensher, D.A., Swait, J.A. (2000). *Stated Choice Methods: Analysis and Application*. Cambridge University Press, Cambridge.
- Love, D.A., Palumbo, M.G., Smith, P.A. (2009). "The trajectory of wealth in retirement". *Journal of Public Economics* 93, 191-208.
- Lucarelli, C., Prince, J., Simon, K. (2009). "The welfare impact of reducing choice in Medicare Part D: a comparison of two regulation strategies". Available at SSRN: <http://ssrn.com/abstract=1217662>
- Lusardi, A., Mitchell, O. S. (2011). "Financial literacy and retirement planning in the United States". *Journal of Pension Economics and Finance* 10, 509-525.
- Lusardi, A., Mitchell, O. S. (2014). "The economic importance of financial literacy: theory and evidence". *Journal of Economic Literature* 52, 5-44.
- Lusardi, A., Mitchell, O. S. (2007). "Baby boomer retirement security: the roles of planning, financial literacy, and housing wealth". *Journal of Monetary Economics* 54, 205-224.
- Madrian, B.C., Shea, D.F. (2001). "The power of suggestion: inertia in 401 (k) participation and savings behavior". *Quarterly Journal of Economics* 116, 1149-1187.
- Maestas, N., Schroeder, M., Goldman, D. (2009). "Price variation in markets with homogeneous goods: the case of Medigap". NBER Working Paper No. 14679.
- Mather, M., Carstensen, L.L. (2005). "Aging and motivated cognition: the positivity effect in attention and memory". *Trends in Cognitive Sciences* 9, 496-502.
- Mazzonna, F., Peracchi, F. (2012). "Ageing, cognitive abilities and retirement". *European Economic Review* 56, 691-710.
- McCall, N., Rice, T., Sangl, J. (1986). "Consumer knowledge of Medicare and supplemental health insurance benefits". *Health Services Research* 20, 633-657.
- McCormack, L.A., Garfinkel, S.A., Hibbard J.H., Norton, E.C., Bayen, U.J. (2001). "Health plan decision making with new Medicare information materials". *Health Services Research* 36, 531-554.
- McFadden, D. (1974). "Conditional logit analysis of qualitative choice behavior". In: Zarembka, P. (Ed.), *Frontiers in Econometrics*. Academic Press, New York.
- McFadden, D. (1986). "The choice theory approach to market research". *Marketing Science* 5, 275-297.
- McFadden, D. (2006). "Free markets and fettered consumers". *American Economic Review* 96, 3-29.
- McFadden, D., Morikawa, T., Ben-Akiva, M. (2002). "Discrete choice models incorporating revealed preferences and psychometric data". In: Franses, P.H., Montgomery, A.L. (Eds.), *Econometric Models in Marketing* vol. 16, Elsevier Science: Oxford.
- Milligan, K. (2005). "Life-cycle asset accumulation and allocation in Canada". *Canadian Journal of Economics* 38, 1057-1106.

- Mitchell, O.S. (1998). "Worker knowledge of pension provisions". *Journal of Labor Economics* 6, 21-39.
- Mitchell, O.S., Moore, J.F. (1998). "Can Americans afford to retire? New evidence on retirement saving adequacy". *Journal of Risk and Insurance* 65, 371–400.
- Mitchell, O.S., Piggott, J., Takayama, N. (2011). *Securing Lifelong Retirement Income: Global Annuity Markets and Policy*. Oxford University Press, Oxford.
- Morrin, M., Inman, J.J., Broniarczyk, S.M., Nenkov, G.Y., Reuter, J. (2012). "Investing for retirement: the moderating effect of fund assortment on the 1/n heuristic". *Journal of Marketing Research* 49, 537-50.
- Mullainathan, S., Noeth, M., Schoar, A. (2012). "The market for financial advice: an audit study". NBER Working Paper No.17929.
- Navarro-Martinez, D., Salisbury, L.C., Lemon, K.N., Stewart, N., Matthews, W.J., Harris, A.J.L. (2011). "Minimum required payment and supplemental information disclosure effects on consumer debt repayment decisions". *Journal of Marketing Research* 48, S60–S77.
- Neuman, P., Cubanski, J. (2009). Medicare Part D update — Lessons learned and unfinished business." *New England Journal of Medicine* 361, 406-414.
- Ooijen, R., Alessie, R., Kalwij, A. (2014). "Saving behavior and portfolio choice after retirement". Panel Paper 42, Network for Studies on Pensions, Aging and Retirement.
- Parente, S.T., Feldman, R., Christianson, J.B. (2004). "Employee choice of consumer-driven health insurance in a multiplan, multiproduct setting". *Health Services Research* 39, 1091-1111.
- Peters, E. (2008). "Numeracy and the perception and communication of risk". *Annals of the New York Academy of Sciences* 1128, 1-7.
- Peters, E., Hess, T., Västfjäll, D., Auman, C. (2007). "Adult age differences in dual information processes: implications for the role of affective and deliberative processes in older adults' decision making". *Perspectives on Psychological Science* 2, 1-23.
- Peters, E., Hibbard, J., Slovic, P., Dieckmann, N. (2007). "Numeracy skill and the communication, comprehension, and use of risk and benefit information". *Health Affairs*, 26, 741-748.
- Polyakova, M. (2015). "Regulation of insurance with adverse selection and switching costs: evidence from Medicare Part D". NBER Working Paper No.21541.
- Poterba, J.M. (2014). "Retirement security in an aging population". *American Economic Review Papers and Proceedings* 104, 1–30.
- Poterba, J. M., Venti, S. F., Wise, D. A. (2011). "The composition and draw-down of wealth in retirement". *Journal of Economic Perspectives* 25, 95-117.
- Rangel, A. (2005). "Comment on passive decisions and potent defaults". In: Wise, D.A., (Ed.), *Analyses in the Economics of Aging*. University of Chicago Press, Chicago.
- Reinholtz, N., de Langhe, B., Fernbach, P.M. (2015). "Almost everyone misunderstands the benefits of diversification". Working paper, Center for Research on Consumer Financial Decision Making, University of Colorado, Boulder.

- Roalf, D.R., Mitchell, S.H., Harbaugh, W.T., Janowsky, J.S. (2011). "Risk, reward, and economic decision making in aging". *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, gbr099.
- Roland, M., Campbell, S. (2014). "Successes and failures of pay for performance in the United Kingdom". *New England Journal of Medicine* 370, 1944-1949.
- Rubaltelli, E., Rubichi, S., Savadori, L., Tedeschi, M., Ferretti, R. (2005). "Numerical information format and investment decisions: implications for the disposition effect and the status quo bias". *Journal of Behavioral Finance* 6, 19-26.
- Sabelhaus, J., Bogdan, M., Holden, S. (2008). "Defined contribution plan distribution choice at retirement: a survey of employees retiring between 2002 and 2007". *Investment Company Institute Research Series*, Investment Company Institute, Washington.
- Saez, E. (2009). "Details matter: the impact of presentation and information on the take-up of financial incentives for retirement saving". *American Economic Journal: Economic Policy* 1, 204-228.
- Samanez-Larkin, G., Kuhnen, G., Yoo, D., Knutson, B. (2010). "Variability in nucleus accumbens activity mediates age-related suboptimal financial risk taking". *Journal of Neuroscience* 30, 1426-1434.
- Skinner, J. (2007). "Are you sure you are saving enough for retirement?" *Journal of Economic Perspectives* 21, 59-80.
- Song, C. (2015). "Financial illiteracy and pension contributions: a field experiment on compound interest in China". Available at SSRN 2580856.
- Spranca M., Kanouse, D.E., Elliott, M., Short, P.F., Farley, D.O., Hays, R.D. (2000). "Do consumer reports of health plan quality affect health plan selection?" *Health Services Research* 35, 933-947.
- Spinnewijn, J. (2016). "Heterogeneity, demand for insurance and adverse selection". Working Paper, London School of Economics.
- Stango, V., Zinman, J. (2009). "Exponential growth bias and household finance". *Journal of Finance* 64, 2807-2849.
- Stern Y. (2002), "What is cognitive reserve? Theory and research application of the reserve concept." *Journal of the International Neuropsychological Society* 8, 448-460.
- Super System Review (2010). *Super System Review - Final Report*, 30 June 2010. Canberra.
- Swait, J., Andrews, R. (2003). "Enriching scanner panel models with choice experiments". *Marketing Science* 22, 442-460.
- Teppa, F., Lafourcade, P. (2013). "Can longevity risk alleviate the annuitization puzzle? Empirical evidence from survey data". Working paper, De Nederlandsche Bank (DNB).
- Tumlinson, A., Bottigheimer, H., Mahoney, P., Stone, E., Hendricks, A. (1997). "Choosing a health plan: What information will consumers use?" *Health Affairs* 16, 229-238.
- Uhrig, J.D., Harris-Kojetin, L., Bann, C., Kuo, T.M., (2006). "Do content and format affect older consumers use of comparative information in a Medicare health plan choice? Results from a controlled experiment". *Medical Care Research and Review* 63, 701-718.

- Van Rooij, M., Lusardi, A., Alessie, R. (2011). "Financial literacy and stock market participation". *Journal of Financial Economics* 101, 449-472.
- Vlaev, I., Chater, N., Stewart, N. (2009). "Dimensionality of risk perception: factors affecting consumer understanding and evaluation of financial risk". *Journal of Behavioral Finance* 10, 158-181.
- Weber, E. (2003). "Who's afraid of a poor old age? Risk perception in risk management decisions". In: Mitchell, O.S., Utkus, S.P. (Eds.), *Pension Design and Structure: New Lessons from Behavioral Finance*. Oxford University Press, Oxford.
- Weber, E.U., Siebenmorgen, N., Weber, M. (2005). "Communicating asset risk: how name recognition and the format of historic volatility information affect risk perception and investment decisions". *Risk Analysis* 25, 597-609.
- Winter, J., Balza, R., Caro, F., Heiss, F., Jun, B., Matzkin, R., McFadden, D. (2006). "Medicare prescription drug coverage: consumer information and preferences". *Proceedings of the National Academy of Sciences of the United States of America* 103, 7929-7934.
- Wu, S., Asher, A., Meyricke, R., Thorp, S. (2015). "Age Pensioner profiles: a longitudinal study of income, assets and decumulation". CEPAR Working Paper 2015/17.
- Wu, S., Stevens, R., Thorp, S. (2015). "Cohort and target age effects on subjective survival probabilities: Implications for models of the retirement phase". *Journal of Economic Dynamics and Control* 55, 39-56.

Table 1: Health Plan Attributes (Minneapolis/St. Paul Market - 1988)

	Basic Medicare	Medicare + Medigap w/o drugs	Medicare + Medigap w/drugs	IPA	HMO
Monthly premium	\$28	\$71 to \$82 (age based)	\$95 to \$109 (age based)	\$53	\$40
Drug Coverage			Yes		Yes
Preventive Care				Yes	Yes
Provider Choice	Yes	Yes	Yes	Yes	
Submit Claims	Yes	Yes	Yes		

Table 2: Stated Attribute Importance Measures
("Tell me if you would . . . to consider a plan")

	"Have to Have"	"Like to Have"	"Doesn't Matter"
Observed Attributes:			
Lowest Premium	23%	59%	18%
Drug Coverage	22%	60%	18%
Preventive Care	32%	55%	13%
Provider Choice:			
Choice of Physician	35%	55%	10%
Choice of Hospital	26%	60%	14%
Low Paperwork	38%	53%	9%
Unobserved Attributes:			
Low Cost-Sharing	31%	60%	9%
Quality:			
Highest Quality	44%	52%	4%
Referral to Specialists	41%	54%	5%
Not Rushed from Hospital	33%	56%	11%

Notes: Each attitude scale was coded: 1= "Doesn't Matter" 2= "Like to Have" 3= "Have to Have". The importance of quality measure was created by summing the three quality related questions and dividing by 3. The importance of provider choice measure was created by summing the two provider choice questions and dividing by 2.

Table 3: Parameter Estimates, Observed Attributes

Observed Attribute:	Intercept	Slope
Premium	.014 (.011)	-.007** (.003)
Drug Coverage	.057 (.912)	.384** (.145)
Preventive Care and No Claims	1.887** (.498)	.766** (.202)
Provider Choice	-.395 (1.081)	1.430** (.489)
Must Submit Claims	Collinear with Preventive Care (Plans with preventive care do not have claims)	-.274** (.130)

Note: The “slope” coefficient must be multiplied by the stated importance weight $S_i = 1, 2, \text{ or } 3$, and the result then added to the intercept to obtain the predicted importance weight for person i . Standard errors are in parenthesis below the estimates. A “**” indicates significance at the 5% level.

Table 4: Mean Utility Weight On Selected Attributes for Different Levels of Stated Importance

	“Doesn’t Matter” S=1	“Like to Have” S=2	“Have to Have” S=3
Drug Coverage	$.057+(1) \cdot (.384) =$.441	$.057+(2) \cdot (.384) =$.825	$.057+(3) \cdot (.384) =$ 1.209
Provider Choice	1.035	2.465	3.895

Table 5: Parameter Estimates, Unobserved Attributes

Unobserved attribute 1: Cost Sharing

Basic Medicare	A_{11}	0
Medigap without Drug Coverage	A_{21}	-.270
Medigap with Drug Coverage	A_{31}	-.355
IPA type HMO	A_{41}	-.414
Group HMO	A_{51}	-.271

Unobserved attribute 2: Quality of Care

Basic Medicare	A_{12}	0
Medigap without Drug Coverage	A_{22}	.269
Medigap with Drug Coverage	A_{32}	.261
IPA type HMO	A_{42}	-.081
Group HMO	A_{52}	.161

Estimates of Equation (3):

$$W_{ip} = 2.688 \cdot S_{ip}^* + v_{ip} \quad p=1 \text{ (cost share), } 2 \text{ (quality)}$$

Note: Unobserved attribute levels for Basic Medicare are normalized to 0 as it is the base alternative. In equation (3), S_{ip}^* is the weight (from 1 to 3) that person i says he/she puts on attribute p , and v_{ip} is “measurement error.”