Economic Data Engineering*

Andrew Caplin†

Draft: May 2016

Abstract

How economics advances will depend on how it contributes to next generation data sets and on how open it is to profound shifts in the landscape of measurement. Scientific and technological advance is driving explosive growth in measurement possibilities. This opens up new vistas for economic theory. This paper outlines the co-evolutionary approach to economics and data that is economic data engineering. It is organized around two basic constructs: beliefs and preferences. It illustrates how data engineering crosses boundaries within and between disciplines.

1 Introduction

How economics advances will depend on how it contributes to next generation data sets. It will depend equally on how open and aware the field is of profound shifts in the landscape of measurement. This paper outlines the co-evolutionary approach to advancing economics and data that is economic data engineering. Need for such engineering derives from both "push" factors

*Draft for *Journal of Economic Literature*. Thanks to Joseph Briggs, Ann Caplin, Olivia Caplin, David Cesarini, Daniel Csaba, Mark Dean, Steve Durlauf, Srijita Ghosh, Mateusz Giezek, Paul Glimcher, Daniel Martin, Filip Matejka, Chris Tonetti, Ruth Wyatt, and Cecilia Wyatt for their valuable contributions.

†Center for Experimental Social Science and Department of Economics, New York University. Email: andrew.caplin@nyu.edu

associated with the limits of standard behavioral data, and "pull" factors related to ongoing scientific and technological advance. This is driving explosive growth in measurement possibilities. By maintaining contact with the expanding borders of feasible measurement, data engineering can open up entire new vistas for economic theory and applied economics.

On the push side, there are fundamental problems of model identification that constrain the progress of our science. One such problem relates to beliefs and preferences. Block and Marschak, 1960, were the first to highlight the difficulty of separating these forces in choice data, when introducing the now ubiquitous random utility model (RUM). Their introduction of stochastic factors into the theory of choice is key for applied work since deterministic theories of choice are generally rejected. While modeling only randomness in utility, Block and Marschak were concerned that, in practice, "perceptibility and desirability" may both impact choice. As they knew, stochastic choice data was first introduced to study humans' limited perceptual abilities (Weber, 1834). In terms of modeling behavior, when an individual picks a particular item from a menu, is it their most preferred item, or may there be preferred alternatives that were not noticed?

Developments in economic theory over the past half century have further highlighted this central identification problem between beliefs and preferences. Imperfect information produces stochastic choice just as do RUMs. So do bounds on rationality. Many modern theories of choice are observationally equivalent in standard choice data. As the spiritual fathers of economic data engineering, Block and Marschak proposed a theoretical program to conceptualize and develop new forms of data to improve identification.

> "Our particular way of defining the class of basic observations and, correspondingly, of the general testable conditions is to some extent arbitrary. The study may thus serve as a start when similar attempts are made under another definition of basic observations." (Block and Marschak [1960], p. 98-99).

While Block and Marschak did not make significant progress on their program, others have done just that, as detailed in the body of the paper. Economists have developed increasingly refined methods of engineering belief data as separate from preferences. Experimental work on belief measurement began with Brier, 1950, with Manski (1990, 2004) playing the decisive role in recent survey-based growth. Ironically, "revealed preference" methods have been used for experimental measurement of preferences isolated from beliefs (Samuelson, 1938, Afriat, 1967, Varian, 1982, and Choi et al. 2007). Survey-based measurements (Barsky et al. 1997, Ameriks et al., 2011, 2015) are now seeding advances in economic theory, illustrating the interaction between progress in measurement and progress in modeling characteristic of data engineering.

While economic theorists may appreciate applied work on data engineering, few have participated. Indeed many may regard it as essentially outside their purview. The advantages of specialization were noted by Adam Smith at the very beginnings of our discipline. We have taken this to heart. Most economic theorists interpret their role as being to model how important underlying forces play out in observable behavior. Econometricians develop methods for estimating these models in suitably rich such data. Applied economists are responsible for gathering the data itself and undertaking model estimation.

With all of its advantages, such specialization makes it hard to enrich observable behavior in a theoretically-guided manner, as Block and Marschak proposed. One illustration of the role of modeling in data engineering derives from rational inattention theory (Sims, 1998 and 2003). As the economic generalization of Weber's perceptual data, **state dependent** stochastic choice data is of particular value (Cover and Thomas, 2012, and Matejka and McKay, 2015). Its precise connection with theory can speed up the interaction with measurement. It can be used to systematically amend estimates of demand in missing markets (Caplin, Leahy, and Matejka, 2016). It can be used to test Bayesian expected utility maximization, the general theory of rational inattention, and more specific variants (Caplin and Martin, 2015, Caplin and Dean, 2015, Caplin, Dean, and Leahy, 2016). As such, it does for attention costs what input-output data does for production costs.

The more general point is that designing enriched data is extremely challenging not only at the practical level, but more importantly at the conceptual level. The models that we develop to explain behavior involve many factors that we treat as unobservable, including not only beliefs and preferences, but also strategies in all contingencies, even those that are unrealized. In the spirit of data engineering, it should be the responsibility of the theorist who is developing a new model to specify the ideal test data set. At best, they would also take measures to ensure that corresponding measurement devices could in principle be developed. This would render their theories more readily testable, and increase feedback from measurement to theory. Ideally, such interplay between measurement and modeling will see both progress in a co-evolutionary manner.

Section 2 of the paper covers engineering of belief data, with preference data covered in section 3. The stress is on best use cases. Section 4 takes up the case of stochastic choice. Section 5 covers data engineering associated with the theory of rational inattention. Research on rational inattention connects back with recent work on updating of beliefs in the face of new information. This is an important subject not only for economists but for those designing educational tests. Data engineering in this area was proposed by De Finnetti, 1965, and Savage, 1971. Section 6 picks up this loose thread and connects it with findings about learning from many different fields, including psychology, neuroscience, and genetics.

Data engineering crosses boundaries within and between disciplines. Within economics, it calls for new hybrid approaches to theory and measurement. In fact it may change not only the boundary between pure and applied work in economics, but across current siloed sub-areas. More broadly, as we increasingly measure interactions between biological, behavioral, and social outcomes, boundaries between larger academic disciplines may gradually erode.

2 Belief Measurement

Qualitative measurement of beliefs has a long history in psychometrics. In a typical survey question in this tradition, respondents place events in such discrete categories as possible, likely, unlikely, etc. As the theory of choice under uncertainty developed, so proposals were made for more quantitative measurement. The early literature was reviewed by Savage, 1971, in making his proposal for a "proper" scoring rule. Pride of place goes to Brier, 1950, who moved weather forecasting in the quantitative direction. As a result of widespread adoption of the quantitative probability measures, the qualitative approach has largely been replaced, even within psychometrics (see Tetlock and Mellers, 2011, and Budescu and Wallsten, 1995).

As the literature has developed, so measurement of probabilities has improved. With regard to experimental measurement, Savage's proper scoring rule was incentive compatible only with risk neutrality. The quadratic scoring rule of Nyarko and Schotter, 2002, is more robust in this respect. A further development is the binary scoring rule of Hossain and Okui, 2013, which requires only linearity in probabilities rather than in terms of dollars. The literature on experimental measurement has mushroomed since this time (see Schotter and Trevino, 2014).

Among the first to contemplate quantitative survey measurement of beliefs concerning future values outcomes was Haavelmo:

> "It is my belief that if we can develop more explicit and a priori convincing economic models in terms of these variables, „, , then ways and means can and will eventually be found to obtain actual measurements of such data.." Haavelmo [1958, p. 357]

The use of survey methods was initiated by Juster, 1966, in the context of future car purchases. He replaced vague questions on intentions with quantitative measures of the likelihood of purchasing, which predicted actual purchases better. After a hiatus, it was the path-breaking contributions of Manski and others (Manski, 1990, Dominitz and Manski, 1996, 1997, and Hurd and McGarry, 1995) that opened the research floodgates. Credit is also due to Richard Suzman, Tom Juster, and Bob Willis. It is only through placement of expectations questions on the Health and Retirement Survey (HRS) that their full value is coming to be appreciated. Corresponding questions are now posed in household panel surveys worldwide.

Manski, 2004, and Hurd, 2009, summarize important findings concerning survey-measured probabilities, including their internal consistency and connection with external reality. For example, Hurd and McGarry, 2002, and Hudomiet and Willis, 2013, show that individuals and groups with higher subjective survival probabilities live longer. Kézdi and Willis, 2011, study the interaction between stock market projections and stock ownership, Hudomiet, Kézdi, and Willis, 2011, study the impact of the financial crisis of 2007-09 on stock market expectations, while Hurd and Rohwedder, 2012, show that differences in these expectations predict the direction of future stock purchases and sales. Wiswall and Zafar, 2015a,b use sequential surveys to understand how provision of objective information on returns to schooling alters understanding. Van der Klaauw, 2012, illustrates incorporation of expectations questions in structural modeling.

As with experiments, the increased use of measured probabilities in surveys in leading to improvement. For example, patterns of error, such as over-use of the 50% focal answer, are producing further innovations in measurement (Fishcoff and Bruine de Bruin, 1999, Bruine de Bruin and Carman, 2012, and Manski and Molinari, 2010). Visual aids are being developed to present probabilistic constructs in as unambiguous a manner as possible (e.g. the "bins and balls" format of Delavande and Rohwedder, 2008). Cognitive interviews have been employed to great effect to improve the design of survey questions on inflation expectations (Armantier et al., 2015).

Measured expectations are of particular importance in exploring phenomena that are hard to explain with the standard expected utility model, but may have many less standard rationalizations. An example relates to investors' apparent unwillingness to realize losses (Odean, 1998). Kuhnen, 2015, investigates this phenomenon experimentally. She identifies a sharp negative impact of realized losses on beliefs, which suggests the value of amendments to standard models of updating. It is as if losses not only grab attention, but may also hijack usual methods of updating. Absent the data on beliefs, this swing to pessimism after a fall in stock prices might have been interpreted as loss aversion, for which little direct evidence was found.

A key reason for measuring expectations is to identify appropriate alternatives to standard models of rational expectations, as stressed by Manski, 2004. Plausible alternative models are increasingly under development. Barberis, Shleifer, and Vishny, 1998, and Barberis et al., 2015, model over-reliance on extrapolation in predicting future trends. Fuster, Laibson, and Mandel, 2010, put forward a theory of "natural" expectations, which represents a mixture between rational and extrapolative expectations. Increasingly, these models are validated at least in part based on survey-based measures of expectations.

A particularly promising market for measuring and modeling expectations is the housing market. Here an extrapolative model of house price dynamics has been developed by Glaeser and Nathanson, 2015. The housing market involves many small and not particularly expert investors having to make bets based on limited understanding of likely future house prices. It is easy to imagine that these expectations may not use all available information optimally, but rather be based on some simple and not entirely implausible theory. In terms of modeling, Glaeser and Nathanson follow Eyster and Rabin, 2010, and model imperfect inference concerning the belief processes of others. The key mistake in such inference is the failure to realize the patterns in prices that will set up if others view the world similarly. This aligns with the low levels of strategic thinking in Nagel, 1995, and Camerer, Ho, and Chong, 2004. While the model is estimated using data on prices, it is inspired in large part by survey-measured expectations documented by Case, Shiller and Thompson, 2012. They find that the pattern of model-implied expectations closely mirrors the corresponding survey findings.

Beliefs about housing returns are inducing particularly interesting new measurements. Malmendier and Nagel, 2011, provide evidence that measuring personal experiences has potentially large effect on beliefs about the future. While their first example is inflation, survey evidence shows the effect to be more general, and to apply also to housing markets (Kuchler and Zafar, 2016). Similarly it is clear that social learning plays a role in the formation of house price expectations. Bailey at al., 2016, present evidence on how experiences of members of an individual's social network impact their personal beliefs. House buyers may seek advice from many parties as they contemplate making a purchase. This makes it of interest to study how well individuals aggregate these different sources of information. In an experimental setting, Enke and Zimmerman, 2003, measure beliefs to explore the possibility of correlation neglect (Akerlof and Shiller (2009)). When given many messages with a particular opinion, they show that subjects tend to update as if they were receiving independent information, even if are told explicity that all message providers derived the information from a common source. They show that this leads to pronounced and predictable price bubbles. Hommes, 2013, has introduced rich experimental markets for the study of such effects. This provides a valuable laboratory for analyzing changes in beliefs and their interaction with price dynamics.

3 Preference Measurement

3.1 Stated Preference Methods

Ben-Akiva, McFadden, and Train, 2015, cover the long history of stated preference measurement. They trace the origins back to a proposal of Thurstone, 1931. This was better received by psychometricians and statisticians than by economists (Luce and Tukey, 1964). The next major impetus in economics related to contingent valuation for public goods. While there remains widespread skepticism about the accuracy of question on how much individuals would pay to maintain biodiversity (Carson, 2012), this literature did lead to valuable refinements of method (Arrow et al, 1993).

In recent years, more economists have perceived the possible value of stated preference methods. The literature was reintegrated into the field by McFadden, 1986. He showed how stated preferences could be analyzed using the tools of discrete choice analysis and the theory of random utility maximization (RUM). Subjects presented with products described in terms of attributes (including price) would be asked to choose their most preferred product in each menu. McFadden showed how choice data from these menus, could be handled in the same way as the real market choice data. Recent applications of these methods include Blass, Lach, and Manski, 2010, on electricity supply, Delavande and Manski, 2015, on political candidates, and Wiswall and Zafar, 2016, on job choice.

A recent paper by Kesternich et al., 2012, provides evidence of the value of stated preference methods in missing market analysis. In the context of a new insurance program, they find that "Hypothetical and real data produce similar estimates of willingness-to-pay (WTP) for insurance plan attributes (Kesternich et al., 2012, p. 3.). They find also that adverse selection is present both in stated demand estimates and in reality so that "hypothetical choice tasks might be used to predict the degree of adverse selection on newly introduced insurance markets as well." This relates to an earlier survey-based finding of Finkelstein and McGarry, 2006, on beneficial selection in the long term care insurance market.

3.2 Revealed Preference and BDM Mechanisms

In 1938, Paul Samuelson introduced his operational appoach to utility theory. The axioms of revealed preference characterize the testable implications of this theory in idealized deterministic choice data. To fully implement his approach at the individual level requires an essentially infinite data set, which led many to see it as an exercise in theory rather than as a guide to mesurement. This started to change when Afriat, 1967, produced his necessary and sufficient conditions for a finite data set to be rationalizable by utility maximization. Varian, 1982, took a critical next step to application. A useable experimental interface was engineered by Choi et al., 2007. There is now an important literature on using these methods to uncover features of individual preference, including how far is behavior from being rationalizable (Dean and Martin, 2015). There is corresponding work at the level of household surveys, including Crawford, 2010.

One recent finding illustrates the new light that revealed preference methods shed on behavioral phenomena. Choi et al., 2014, place their measure of consistency with revealed preference on a large scale household survey. They find large and consistent individual differences. They conclude that consistency with utility maximization under laboratory conditions captures decision-making ability that applies across domains and influences important real-world outcomes. It is striking that economic theory produces such a potentially valuable new tool for exploring individual differences. Falk et al., 2015, develop survey instruments to measure many economically important individual differences on a worldwide basis.

In a typical revealed preference experiment, there is no direct incentive for accuracy. Yet within the standard experimental approach, such incentives are very important. The pioneers in the design of incentive compatible measurement were Becker, De Groot, and Marschak, 1963. With regard to risk aversion, the Holt-Laury instrument has been widely used in the laboratory, with refinements continuing to this day (Holt and Laury, 1972).

3.3 Survey Measures

The amounts of money involved in experiments are too small to trigger risk aversion relevant to life cycle spending. For that reason, Barsky, Juster, Kimball, and Shapiro, 1997, constructed a stated preference question that placed enough wealth on the line to introduce significant wealth swings. It involved a switch of job with a potentially large change in income. With the advantage again of being able to place these on the HRS, this form of question is now widely used and related to portfolio choice. The investigation of survey errors has been initiated by Kimball, Sahm, and Shapiro, 2008. A key insight from these methods relates to the profound underlying heterogeneity that is revealed in preference parameters.

There is particular intricacy in posing stated preference questions relevant to large scale changes in wealth, health, age, and other important state variables. One key aspect of this is that all respondents are at different stages in the life cycle and have other state variables that are personal. In the context of any sequential decision problem, such differences in circumstance may systematically change the mapping from survey response to model parameters. For example, a change in future income will have a very different impact on someone who intends to retire shortly than on a respondent who is just entering the labor force. This may interact in a complex manner with preference parameters. This makes it of value to pose questions that, to the maximum extent possible, put respondents into analogous situations. Interestingly, dynamic programming suggests that all are conditioning current decisions on the behaviors that they would choose in later contingencies. Noting this, Ameriks et al., 2011, introduced "strategic survey questions" (SSQs), which are stated preference questions designed to place respondents in a common future contingency, which, should it occur, would be particularly revealing about their underlying motivations.

The paradigm application concerns the finding that wealthy older households do not spend wealth down in retirement as the simple life cycle model would predict. While all current research assigns responsibility to some combination of bequest motives and precautionary savings driven by high health-related costs, including long-term care, their relative importance has been much debated. The identification problem is hard to avoid, since wealth is fungible (Dynan, Skinner, and Zeldes, 2002). In terms of motives for retaining wealth, a revealing contingency would arise if the respondent faced an explicit insurance decision that restricted its future use. Ameriks et al., 2015, develop a rich and rigorously constructed set of SSQs for estimation of a state of the art model of late in life spending behavior. Several of these questions are stated preference questions about Arrow securities that deliver resources in key contingencies, rather than about different goods at a store. The method is more general than this, and can in principle be used to probe contingent behavior in many individual, social setting, and market settings. Brown, Goda, and McGarry, 2015, use SSQs to show that individuals value wealth more when facing physical rather than mental disabilities requiring long-term care. This is a distinction that would be essentially impossible to make in standard behavioral data.

3.4 Co-Evolution

In qualitative terms, a key finding in Ameriks et al., 2015, is that subjects allocate more to long-term care needs than to end of life bequests. The recovered model parameters liberate analysis of demand in missing markets. Indeed, one of the main reasons for structural modeling is to conduct counter-factual analysis. Ameriks et al, 2016a, use their model in precisely this manner to estimate demand for an idealized form of long-term care insurance. Given the high precautionary motives that they identify, it is not surprising that they estimate very high interest in this product.

In addition to providing model-based estimates, Ameriks et al., 2016a, follow the stated preference approach in the spirit of Kesternich et al., 2012. Again, they identify high demand for idealized insurance. Yet there is a large difference. The model-based estimates are far higher than those based on direct stated preference. This is doubly true for annuities.

The gap between model-implied and stated demand serves as a form of model specification test. Unlike the model-based estimate, the stated preference question is not tied to any particular formulation of the motives for retaining wealth. Hence the finding of relatively lower stated than model-implied demand raises the possibility that the estimated model is incomplete in important respects. This has stimulated both a new data gathering effort and a new model to capture possible missing motives in relation to the family (Ameriks et al., 2016b). Not all such motives appear to be captured in standard models of the bequest as a warm glow (De Nardi, 2004). This form of joint progress in theory, data, and estimation would not have been possible without the consciously engineered new data on preferences.

Some next steps in the care-related agenda involve digging more deeply into inter-generational interactions (Caplin, Luo, and McGarry, 2016). Factually, what is needed is a fuller accounting of the flow of resources, both money and care-giving between generations. This also involves measuring expectations concerning the future and their potential impact on earlier decisions, such the employment and labor supply decisions of children. SSQs can be posed to understand how compromises would be reached when parents and offspring have different interests. There is also the direct question of who has effective control of decision making as adult mental competence declines. Guistinell, 2012, pioneered in the development of corresponding questions in relation to educational decisions. More broadly, it is hard to know how to make progress in the area of family-based decision making without asking well directed questions. There are countless theories of interaction. Engineering data to tease them apart is set to become an increasingly active research frontier.

4 An Identification Problem

4.1 Random Utility Models

Econometric analysis of missing markets started in earnest with the logit model of McFadden, 1974. In a typical setting, agents have preferences over available goods that depend on such observables as age and wealth. There is also unobserved heterogeneity in preferences. If the utility contribution of the observable features of option $a \in A$ is $U(a)$ and the unobservable component of utility is extreme-value distributed, McFadden shows that its market share is,

$$M(a) = \frac{e^{\beta U(a)}}{\sum_{a' \in A} e^{\beta U(a')}}.$$

Here $\beta > 0$ scales the unobservable relative to observable preference factors.

Given data on market share, this formula allows the econometrician to back out estimates of the utility of all available commodities to each type of consumer. The model has been applied far and wide, from such important decisions as school selection and choice of partner, to such every day examples as choice of mode of transport and what meal to buy in a restaurant. One can enrich the model in many ways to structure the population heterogeneity in preferences and how it is related to product characteristics. For example, Berry, Levinsohn, and Pakes, 1995, model the impact of characteristics and prices allowing for type specific utility weights on characteristics and for a good specific price elasticity of demand.

While flexible in many ways, currently estimated models typically take a strong stand on the underlying source of stochasticity in choice. Since the logit model is a RUM, choice probabilities depend only on preferences over the goods. Information is assumed to be perfect. Moreover, in application it is the econometrician who defines these goods. This may be reasonable for the simplest of goods. However credibility is more questionable for goods that are complex and hard to define objectively.

To appreciate the importance of this issue, consider again the analysis of potential interest in ideal long-term care insurance. Ameriks et al., 2016a develop both the stated preference questions and the SSQs in relation to an ideal product. One reason for this is that current products are notoriously poorly designed, involving limited coverage, the possibility of large future changes in premia, possible bankruptcy of the insurer, and a potentially adversarial claims process. A second reason is that the complexity of the current market is disturbing even to professionals (Society of Actuaries, 2014). Demand depends on unmeasured beliefs about insurer solvency, future premium increases, and the difficulty of the claims process. Absent this data, it is hard to confidently infer preference parameters from low demand for current products, as do Koijen, van Niewerburgh, and Yogo, 2015. As work in this area advances, so richer hybrid models and estimated using yet richer data on relevant beliefs and preferences.

The point is general. It is hard to know how objectively to define a particular school, job, partner, or location. If the econometrician finds this difficult to assess, what are the odds that all market participants saw the goods in just the prescibed manner? This again raises the challenge of how to separate out beliefs from preferences. Interestingly, this point was first noted by Block and Marschak, 1960, when introducing RUMs into economics.

4.2 Random Perception and Stochastic Choice

Block and Marschak were looking to characterize RUMs in idealized stochastic choice data. They also looked to characterize the additional behavioral restrictions associated with special models, such as the model of Luce, 1958, with its independence properties. Falmange, 1971, made key contributions. After a lull, RUMs are undergoing something of a revival following the work on expected utility of Gul and Pesendorfer, 2006, with Apesteguia and Ballester, 2015, contributing a new formulation in which preference parameters rather than choices are subject to random variation.

There is a telling difference between the deterministic choice data of Samuelson, 1938, and stochastic choice data. Samuelson was concerned with choice from budget sets at different levels of income and at different prices. This data set is firmly rooted in the economic tradition. In contrast, stochastic choice data derived from the psychometric tradition. The Weber-Fechner laws of psychophysics highlight the distinction between the objective differences between two stimuli, as known to the experimenter, and the subjectively perceived difference. The first formal model of stochastic choice data was that of Thurstone, 1927, who modeled comparisons between stimuli as based on perception as opposed to reality. Even Luce, (1956, 1958) referred to choice probabilities as defining a "discrimination structure". To this day, the final step in psychometric models of choice is to add stochasticity through application of a "soft max" function. Given real numbers $V(a)$ for $a \in A$, the frequency with which an experimental subject judges n to be the largest is modeled as,

$$\sigma(a) = \frac{e^{\beta V(a)}}{\sum_{a' \in A} e^{\beta V(a')}};$$

with $\beta > 0$ scaling the subject's ability to discriminate between values.

4.3 The Identification Problem

The soft-max and the logit form are essentially identical. One and the same function can be used to summarize choice probabilities as resulting from differences in utility, differences in perception, or both. Block and Marschak were acutely aware of their data set's origins and of the implied identification problem.

> "In particular, our operational approach seems to be unable to handle the following distinction that appears natural on grounds of common sense and may be important for predictions. If out of the pair $F = (a, b)$ of desirable objects a man chooses sometimes a and sometimes b, our introspection tells us that we may ascribe this to either or both of two different "causes":

1. He may have difficulty in perceiving all the relevant characteristics of the objects...

2. Even if he knew exactly the differerences of the characteristics of the two objects, he might find them almost equally desirable ... and he will vacillate as a result.

> To distentangle the two "causes" – call them perceptibility and desirability (anticipated "satisfaction") – may be important if one wants to predict how people will act if perceptibility is kept constant while desirability varies, or vice versa." (Block and Marschak, 1960, p. 99).

4.4 Imperfect Information and Stochastic Choice

It may seem that the disagreement is between economics and psychology, but the issue is far deeper. Imperfect perception of available options is central to economics. While economists do not typically see themselves as modeling imperfect perception, in practice they do under its pseudonym, imperfect information. What is learned in all models of search and signal processing is random. This produces randomness in choice, as do models of boundedly rational choice (Simon, 1955, Gabaix et al., 2006, Salant and Rubinstein, 2008, Manzini and Mariotti, 2007, 2014, Masatogliu, Nakajima, and Ozbay, 2011, Caplin and Dean, 2011, Caplin, Dean, and Martin, 2011, and Gabaix, 2014).

The identification problem is particularly acute given that almost all learning is costly. As first formalized in the theory of search (Stigler, 1961), this creates a link between comprehension and utility. "Perceptibility and desirability" are not independent. This raises the importance of teasing them apart. Yet the interdependence also suggests possible ways forward, using theory to discover structure in this relationship.

5 Rational Inattention and Data Engineering

5.1 The Block-Marschak Proposal

In addition to pinpointing the basic identification problem, Block and Marschak dedicated a section of their paper to data enrichment:

> "Our particular way of defining the class of basic observations and, correspondingly, of the general testable conditions is to some extent arbitrary.
>
> ..by using a particular demarcation of the class of directly testable conditions (the one most closely corresponding to the nature of economic observations), we are able to carry out a reasonably complete analysis of the relevant logical relations. The study may thus serve as a start when similar attempts are made under another definition of basic observations." Block and Marschak, 1960, p. 98-99).

Despite the centrality of this identification problem, progress in defining new basic observations has been limited. It is currently picking up again. In this paper we stress progress associated with the theory of rational inattention due to Sims (1998, 2003). Intriguingly, while Marschak did not formulate the model, he felt that Shannon's work was more important that economists had realized (Marschak, 1974). It might not have surprised him that it provides such important pointers to data engineers.

5.2 Rational Inattention in a Market

If one envisions the DM's internalization of information as guided at least in part by the potential uses of this information, it is natural to connect them in a unified modeling framework. This is the role of rational inattention theory. Matejka and McKay, 2015, show that this theory produces randomness in choice that is powerfully related to, yet distinct from, the pattern in the standard logit model. The noise in choice is driven by noise in signals that the agent receives. The only difference from the standard form of the logit model is the inclusion of unconditional choice probabilities that are based on prior beliefs.

Caplin, Leahy, and Matejka, 2016, extend the model to allow for social learning. In this context they show how rational inattention impacts discrete choice in a market. Each period entering agents select among a fixed set of available alternatives. As in Caminal and Vives, 1996, they freely observe past market shares of all alternatives. The innovation is that they can also choose to acquire additional private information about them. It is here that rational inattention enters, since additional private learning is assumed to be costly. This cost is some multiple $\lambda > 0$ of the Shannon mutual information between prior and posterior beliefs, as in Woodford, 2009. The incorporation of social learning is important in many applications, particularly when decisions are complex. Cases in which it has been identified range all the way from adoption of production technology (Foster and Rosenzweig, 1995, Munshi, 2004) to choice of retirement savings plan (Duflo and Saez, 2003), health insurance plans (Sorensen 2006), and choice in restaurants (Cai, Chen, and Fang, 2009).

In Caplin, Leahy, and Matejka, market participants start out with a common prior over the (finite) space of types, $\omega \in \Omega$. The effect of past observation is to drive all new agents to common updated beliefs over the possible preference types. As in Matejka and McKay, 2015, choice probabilities depend on a mapping from this intial belief and the available options $a \in A$ into unconditional choice probabilities, $P(a)$. These unconditional probabilities are sufficient to characterize each type's stochastic choice as,

$$P(a|\omega) = \frac{P(a)e^{U(a,\omega)/\lambda}}{\sum_{a' \in A} P(a')e^{U(a',\omega)/\lambda}},$$

where $U(a, \omega)$ is the expected utility of option a to a consumer of type ω. Necessary and sufficient conditions characterizing optimal unconditional choice probabilities are in Caplin, Dean, and Leahy, 2016a, who note that many options that are unlikely to be best may be unchosen.

From the viewpoint of inference, the key result is that market shares converge. What this does is to render the model operational, since these market shares define the priors of new entrants concerning how likely they are to be of each preference type. Observed market shares capture all relevant information about the distribution of types. By definition, unchosen options are not represented in market data.

5.3 Cross-Market Restrictions

Caplin, Leahy, and Matejka show that long run market shares weight together social and private learning in a direct, simple, and entirely reasonable manner. The more utility a particular option provides, the more likely it is that the agent selects it, and this dependence is stronger for lower costs of information, λ. Yet the observed market shares have a systematic influence in distorting choice. This implies that there is a potential bias in inference of utility parameters if the information frictions are present, but neglected. Since high market share attracts demand from those who are inattentive, the effects of characteristics associated with high market share options will be biased upward and those with low market shares are biased downward. In an industrial organization setting, this suggests biases in elasticity estimates in prevalent approaches.

The theory suggests methods of inference that involve looking across markets. Separation of private and social costs is possible if one observes the same market under various different conditions. The change may be as simple as changing a price of one of the goods, or as comprehensive as studying demand for a new set of goods. The necessary and sufficient conditions allow one to identify precisely those products that will be chosen in equilibrium. The broader method of using cross-market restrictions for purposes of identification in discrete choice settings was pioneered by Heckman and Honore, 1990.

5.4 State Dependent Stochastic Choice Data

Rational inattention produces a non-standard information asymmetry. An outside observer with access to suitably rich data on market shares may be better able to understand preferences than are decision makers themselves. The reason for this is that individual choice probabilities reflect some mix of personal preferences and inferences from the broader distribution of population preferences as reflected in market shares. Aggregating across similar types according to some additional observables beyond the crude market aggregates that impact prior beliefs can therefore enrich inference. The most detailed demand data that could be observed conditions precisely on consumer type, $\omega \in \Omega$. In the model of Caplin, Leahy, and Matejka, 2016, observing this data would allow the econometrician in principle to identify the expected utility function of each type.

The value of observing the relationship between demand and truths that are not necessarily available to market participants is general. In the general model of rational inattention, the state dependent stochastic choice (SDSC) data is ideal. This specifies state dependent probabilities $P(a|\omega)$ for all actions and states. The precise state in question depends on the application. Chetty, Looney, and Kroft [2009] analyze the impact on demand of inclusion or exclusion of sales' tax in stated price. One can formulate this as an observation of stochastic demand conditioning on a state that defines the translation of the stated price to the final sales price of the item. This is known to the store and/or experimenter involved, but may not be fully internalized by inattentive purchasers. The precise subject of analysis is the extent to which this external reality appears to be understood by those who enter the store, which can readily be stated in the language of rational inattention. The mapping from experiment to model requires the analyst to be explicit about the prior beliefs of purchasers. One possibility in this case is that the prior measures the proportion of the time sales' tax is included in the stated price of goods in the corresponding field experiment. In a similar manner, Martin, 2016, uses scanner data on grocery purchases to test various model of inattention to prices. In this example, the prior is the empirical distribution of prices and SDSC is stochastic choice of bundles at various prices, which is the standard stochastic demand function.

SDSC has much to recommend it in terms of the desiderata for data engineering laid out by Block and Marschak. One of these was that it should correspond as closely as possible "to the nature of economic observations." One can view SDSC data as very close to this ideal in light of the examples above. Moreover, while strong assumptions are needed to define market observations in a manner that matches the ideal of SDSC, this is easy in an experimental settings. Indeed the perceptual data sets of Weber and ensuing psychometricians studying discrimination are of precisely this form. It is reasonable to assume, as Weber did implicitly, that the prior that the heavier weight is in either hand is 0.5. In essence, SDSC data is the general version of this standard psychometric data set for an arbitrary number of options, an arbitrary incentive scheme, and an arbitrary prior belief on how likely is each option to be of each possible type. Corresponding data has now been gathered in a number of economic experiments (Dean and Neligh, 2016).

The second key requirement for "basic observations" laid out by Block and Marschak is the ability to use them to characterize theories of choice. SDSC stands out as unique in this regard. Caplin and Martin, 2015, characterize Bayesian expected utility maximization in this data set for any given decision problem. Caplin and Dean, 2015 characterize an entirely general model of optimal costly information acquistion by looking across decision problems. Note that this makes no restriction on functional form, and in that sense is equivalent to the most general characterizations of utility maximization. Caplin, Dean, and Leahy, 2016b, characterize behavioral data associated with special cost functions, including the now ubiquitous Shannon cost function, which is remarkable in terms of its computational simplicity. In terms of behavior, it involves a number of qualitative restrictions. As our understanding of these behavioral properties deepens, so less restrictive cost functions will be developed. Indeed Woodford, 2014, uses behavioral patterns in such data to cast doubt on the Shannon model. It will be of particular interest to characterize cost functions to match such intuitive properties of attentional effort as complementary acts of learning, simplicity of comparison, and returns to attentional effort. As the characterizations are discovered, attentional cost functions may become as varied as technological cost functions and utility functions, and be equally essential to economic analysis.

As with standard choice data, SDSC data is of interest even when the theory that it is initially designed to test blatantly fails the specification test. To use it as a basis for data engineering is not to believe in it, but to see its value as an organizing system for uncovering where to change and where to move forward. In introducing the revealed preference approach, Samuelson [1938] made clear his view that the choice data was more fundamental than the model of utility maximization, which, if false, should be rejected in favor of models that more accurately characterize observed patterns of behavior (Dixit, 2012). Hence SDSC data may be of interest when the decision maker is not rationally inattentive, and even when Bayes' rule is not adhered to. Of particular interest are behavioral deviations from rational expectations and from Bayesian updating.

6 Testing and Learning

Experimental protocols for capturing SDSC are simple and general. They involve discrimination tasks. One seeks a setting in which a large number of subjects are given a certain amount of time to pick among optons the rewards to which are initally unknown. With attentional effort, the subjects can improve their understanding and hence the likely quality of their choices. As in the Weber experiments on weight discrimination, there must be some natural method for assigning prior beliefs over the answers. Moreover the choices that are made must reflect to some extent the actual discrimination ability of the subjects.

A little reflection suggests an analogy between these experimental desiderata and a context with which we are all familiar: multiple choice tests. In this section we explore the possible value of data engineering in the areas of testing and of teaching. We identify many forms of data that may be relevant in this area. This is ultimately an area of great policy relevance. Improvements in measurement have the potential to lead to more personalized and effective teaching and learning. Not surprisingly, these issues are already of interest to scholars in many different disciplines. This breadth of engagement is reflected by the many forms of data that may be of interest in this area of application.

6.1 Item Response Theory

A standard test involves a group of students taking the same exam, the score is which is used as a basis for inferring skill. Typically, the score is computed by adding up all correct answers, with a fractional penalty for incorrect answers. While ubiquitous, it is hard to rationalize this simple approach as correctly identifying any underlying skill. The only theory that supports this process in any way is “item response theory”, which posits existence of a single latent skill trait that correspondingly shifts the probaility of correctly answering any test question.

There is no strong reason to believe that this theory is adequate to capture actual exam performance in tests, let alone later real world behaviors that are dependent on how much the corresponding skill has been internalized. The limits of such a one-dimensional viewpoint are stressed by Borghans, Duckworth, and Heckman et al., 2008. Indeed there is good reason to believe that current protocols allow many social and psychological differences to impact the skill rating. Gender effects are particularly well established (Gneezy, Niederle, and Rustichini, 2003, Niederle and Vesterlund, 2010).

Despite its obvious importance, the issue of how to systematically improve measurement has not received due attention. It certainly involves enriching the data that is used to judge skill, but how? This is particularly challenging in the modern era in which one can literally get thousands of applicants for scarce skill-related positions. Some simple scales must be used to convey aptitude, but what methods can be used to generate and validate these scales?

There are powerful analogies between the factors that are pushing economic theorists into considering non-standard forms of data and the factors that make inference hard in the educational arena. In both cases, one is hoping to use a relatively limited data set to infer something fundamental about the decision maker. In both cases, the assumptions that justify simple methods of inference are dubious. Given that economic theorists and educators face an analogous challenges in this regard, it may be of particular value for them to join forces in the design of data enrichments.

6.2 Beliefs and Utility Revisited

Even in 1965, multiple choice tests were ubiquitous. As a pioneer of subjective probability theory, de Finetti, 1965, argued that the standard method of grading tests was inadequate. He outlined many methods of eliciting richer information on subjective beliefs. His goal was clearly stated in his abstract:

> "It is argued that where a person is uncertain as to the true answer in a multiple choice question, he should be encouraged to express his partial knowledge in terms of the subjective probability he attaches to each alternative being correct. A variety of answering techniques are examined, together with ways of scoring them, to determine how far they provide an adequate appraisal of the subjective probabilities." De Finetti, 1965, p. 87

The proposal of De Finetti connects in a direct manner with economic research on belief elicitation. Yet while this is a huge advance over the standard all-or-nothing response method, additional enrichments are of value. De Finetti was concerned only with eliciting subjective beliefs about how likely is each answer to be correct. In an exam, confidence is not all that matters. Another important feature of examinees is their realism. Some may systematically hold over-optimistic beliefs, others less so (or even the reverse). If probabilities are successfully elicited, what points should one give to a respondent who is confidently wrong relative to one who blatantly has no idea?

Another limitation is that the proposed methods relate to a single question. In a test with multiple questions, the belief elicitation methods that de Finetti posited tell only part of the story. Savage, 1971, considered students answering multiple questions in a winner-take-all exam. He noted that a student seeking to maximize the probability of getting the highest score would behave very differently from an individual looking to maximize expected score. Another obvious confound arises when there is a "guessing" penalty for incorrect answers. In such cases it may be dominant to not answer a question even if the respondent is almost sure of the answer. This depends on how the small probability of being penalized impacts the probability of failing the exam. Again, there is an inference challenge. How can one know if a question was not answered because of a lack of confidence or to avoid subjectively assessed possible penalty? Making this particularly worthy of study is evidence of differential willingness to guess by gender (Baldiga, 2013).

Response strategies are particularly pertinent for exams that rely on a set of techniques that are reapplied in different settings. In such cases there are likely to be correlated beliefs. If there are two conceivable methods of answering a question, then use of the same method on both questions will produce positively correlated probabilities, while switching will produce negatively correlated probabilities. From the viewpoint of a standard test, this sets up an incentive to plunge in cases in which utility is linear, yet possibly to diversify in cases of risk aversion. For example, if the examinee believes that (A, A) is the most likely pair of answers to two related questions with (B, B) the other possibility, then it is rational if in need of only one point to pass to answer (A, B). A third confound relates to the private utility of knowledge and test success. Current methods are based on the assumption that students make an honest effort to answer each question as well as possible. In practice, some may hate a subject or find it of little value, making inference of skill yet more complex. Where is this revealed?

6.3 Rational Inattention and Test Design

The above points strongly to the limitations of standard "single response" multiple choice tests and standard methods of grading them. It points to the need both for enriched response schemes and enriched methods of inference. At this nascent stage of the research, the stress is on the process of data engineering rather than predicting the forms of measurement that will ultimately be needed.

One data set that may be of interest to explore is associated with the theory of rational inattention and costly information acquisition. As noted at the outset, the test format does involve the transformation of an uninformed prior through the input of attentional effort into a more informed posterior. The manner in which attentional effort is guided impacts the likely outcome of the test, which in turn impacts the expected utility of the respondent. How well test takers are able to transform prior to posterior depends on some personal costs of improved discrimination of the answer most likely to be correct. This immediately suggests that the kinds of data sets gathered to test theories of rational inattention may be of interest in the exam context.

In the context of a test, a particularly interesting data set involves giving a test with a wide range of very different grading schemes. In theory this can be extremely revealing. Consider for example the case of two subjects who both get 60% of questions right and 40% wrong in a standard test in which there is no penalty for incorrect answers. Suppose further that they are trully different in that one of them is falsely confident, believing if certain that they got all answers correct. Suppose that the other is fully self aware and confidently identified 50% of answers, while outright guessing on the remaining 50% of the questions. Now suppose that they are told that this same exam may be graded with a different scheme that involves a full point penalty for errors. In this case, the overconfident individual will be revealed by the fall to a 20%

grade, while the realist will get a 50% grade. A full characterization of what such a rich data set can reveal may be of interest.

From the viewpoint of economic modeling, the connection with rational inattention theory may be of value for the earlier question of how much to learn leading up to the test. Once the test protocol is understood, it impacts the costs and benefits of effort dedicated to learning. This depends not only on the future value of being better prepared in a particular dimension, but also on the anticipated nature of the test and how well it will reflect the actual level of skill. It depends also on a personal assessment of the production function for the corresponding skill as a function of the student's input of resources in terms of time, effort, etc. In technical terms, this form of investment decision can be seen as designing a mental structure to operate as an information process in the sense of Blackwell and Shannon. In this sense, one may model the process of learning using much the same apparatus as described above for modeling attentional effort more broadly.

6.4 Time Use and The Drift-Diffusion Model

The above only scratches the surface of the decision problem that is taking a multiple choice test. Consider the time constraint. While there is a long history in psychometrics of measuring decision time, the literature in economics is of more recent vintage (e.g.Wilcox, 1993, Kocher and Sutter, 2006, Rubinstein, 2007, Chabris et al., 2009, Spilliopoulos and Ortman, 2014, Geng, 2015, and Agranov, Caplin, and Tergiman, 2015). In the case of the exam, the key issue is the need to set stopping times even when unsure of the correct answer to a question, with the presumption that spending more time would to some extent improve resolution. This connects exams with another key psychometric model in the shape of the drift-diffusion model (Ratcliff, 1978, and Ratcliff et al., 2016). Variants of this model are increasingly making their way into economics (see Fehr and Rangel, 2011, Krajbich and Rangel, 2011, Krajbich et al., 2012, and Fudenberg, Strack, and Strzalecki, 2015, Caplin and Martin, 2016). In the classical experiment, a prize is available in one of two locations, and there is a flow of evidence indicative of which box contains the prize. The decision maker decides when to stop the flow of evidence by picking a location, and which location to pick. The data set that is produced is precisely SDSC data, with the particular feature that included among the actions is the possibility of delaying choice.

That score depends on time allocation strategy hugely complicates taking the test and impacts what can be inferred from the score. Some may skip difficult questions and come back to them at the end. Others may be more rigid in their application of the linear order. Others may use a hybrid strategy of exploring questions superficially to seek evidence on how easy they are likely to be, before deciding whether to continue or to move on. With regard to data engineering, what becomes important is to explore how changes in the order of questions and the timing protocols impact comprehension as revealed in the structure of responses. The fact that this depends also on the grading scheme makes the challenge of modeling and measurement more profound still.

6.5 Inter-Disciplinary Links

The optimal quitting time problem in test taking is to all intents and purposes insoluble in real time. Hence students may have developed broadly applicable rules of thumb in test taking, that may be more or less well suited to each particular test they face. The evaluation of informational rules of thumb is an increasingly important area within psychometric research (e.g. Gureckis and Markant, 2012). Intriguingly, among the behaviors that have been identified is excessive interest in learning more for the sake of it, even when the reward structure suggests that one should be more targeted. Designing models and corresponding measurements in this are suggests a strongly linked research agenda.

The link with psychometric research is but one small part of a far broader set of inter-disciplinary links that are involved in modeling, measuring, and ultimately enhancing, the educational process. The connection with the drift-diffusion model is a link not only with psychometrics, but also with neurobiology, given the findings of Shadlen and Newsome, 2001, relating to the neural implementation of the model. Stepping back further, there is a rich neurbiological literature on methods of learning, much of it associated with crucial findings on reward prediction errors and the dopamine system (e.g. Schultz, Dayan, and Montague, 1997, Bayer and Glimcher, 2005, Daw and Doya, 2006, Dayan and Daw, 2007, Caplin et al., 2010). Applications to finance are particularly exciting (e.g. Frydman et al., 2014). Camerer, Loewenstein, and Prelec, 2005, Glimcher, 2011, Fehr and Rangel, 2011, and Bernheim 2013, offer broader perspectives on how the advent of neuroeconomics will impact economic theory.

The above shows that data engineering crosses traditional disciplinary boundaries. Examples of this form will become common-place as we measure interactions between biological, behavioral, and social outcomes over the life cycle (Azmak et al., 2015). In the case of education, Rietveld et al. 2014, have recently identified a set of single nucleotide polymorphisms (SNPs) associated with years of schooling. The findings appear robust across generations and geographies, but there is at present no way to understand pathways. In richer data, it will become possible to measure and model cognitive and other pathways. Thereby hangs another tale.

7 Concluding Remarks

This article highlights the importance of next generation data sets for the evolution of economic theory.

8 Bibliography

References

Afriat, Sidney. 1967. "The Construction of Utility Functions from Expenditure Data." *International Economic Review*, Vol. 8, No. 1, pp. 67-77.

Agranov, Marina, Andrew Caplin, and Chloe Tergiman. 2015. Naive Play and the Process of Choice in Guessing Games.*Journal of the Economic Science Association*: 1.2: 146-157

Akerlof, G.A. and R.J. Shiller. 2009. *Animal Spirits: How Human Psychology Drives the Economy, and Why it Matters for Global Capitalism*. Princeton University Press.

Ameriks, John, Joseph Briggs, Andrew Caplin, Matthew Shapiro, and Christopher Tonetti. 2015. "Long-Term Care Utility and Late in Life Saving." National Bureau of Economic Research, Working Paper 20973.

Ameriks, John, Joseph Briggs, Andrew Caplin, Matthew Shapiro, and Christopher Tonetti. 2016a. "Late-in-Life Risks and the Under-Insurance Puzzle." Mimeo, New York University.

Ameriks, John, Joseph Briggs, Andrew Caplin, Mi Luo, Matthew Shapiro, and Christopher Tonetti. 2016b. "Inter-Generational Transfers and Precautionary Saving."

Ameriks, John, Andrew Caplin, Steven Laufer, and Stijn Van Niewerburgh. 2011. "The Joy of Giving or Assisted Living: Using Strategic Surveys to separate Bequest and Precautionary Motives." *Journal of Finance* 66.2: 519-561.

Apesteguia, Jose, and Miguel A. Ballester. 2015 "Monotone stochastic choice models: The case of risk and time preferences." Mimeo.

Armantier, O., Bruine de Bruin, W., Topa, G., Klaauw, W., & Zafar, B. 2015. "Inflation expectations and behavior: Do survey respondents act on their beliefs?" *International Economic Revie*w, 56(2), 505-536.

Arrow, K., R. Solow, P. R. Portney, E. E. Learner, R. Radner and H. Schuman. 1993. *Report of the NOAA Panel on Contingent Valuation*. Federal Register, 58, 4601–4614.

Azmak, Okan, et al. "Using Big data to understand the human condition: The Kavli HUMAN project." Big data 3.3 (2015): 173-188.

Baldiga, K.A. 2013. "Gender Differences in Willingness to Guess. Management Science." Vol. 60, Iss. 2, pp. 434 – 448.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer. 2015. "X-CAPM: An Extrapolative Capital Asset Pricing Model." *Journal of Financial Economics* 115, 1-24.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. 1998. "A Model of Investor Sentiment." *Journal of Financial Economics* 49, 307-343.

Barsky, R. B, F. T. Juster, M. S. Kimball, & M. D. Shapiro, 1997. "Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study." *Quarterly Journal of Economics* 112.2: 537-579.

Bayer, Hannah M., and Paul W. Glimcher. 2005. "Midbrain dopamine neurons encode a quantitative reward prediction error signal." *Neuron* 47.1: 129-141.

Becker, G. M., M. H. DeGroot, and J. Marschak, 1963. "Stochastic models of choice behavior." *Behavioral Science*, 8(1), 41–55.

Ben-Akiva, Moshe, Daniel McFadden, and Kenneth Train. 2015. "Foundations of Stated Preference Eliciation of Consumer Behavior and Choice-Based Conjoint Analysis." Mimeo.

Bernheim, B. Douglas. 2013. "The psychology and neurobiology of judgment and decision making: What's in it for economists?" Chapter 9. *Neuroeconomics: Decision Making and the Brain.* Glimcher, Paul W., and Ernst Fehr, eds. Academic Press, 2013.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile prices in market equilibrium." *Econometrica*: 841-890.

Blass, Asher A., Saul Lach, and Charles F. Manski. 2010. "Using elicited choice probabilities to estimate random utility models: preferences for electricity reliability." *International Economic Review* 51.2: 421-440.

Block, Henry David, and Jacob Marschak. 1960. "Random orderings and stochastic theories of responses." *Contributions to Probability and Statistics* 2: 97-132.

Borghans, L., A. Duckworth, and J. Heckman. 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Resources* 43.4: 972-1059.

Brier, Glenn W. 1950. "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review* 78: 1–3.

Brown, Jeffrey R., Gopi Shah Goda, and Kathleen McGarry. 2015. "Heterogeneity in State-Dependent Utility: Evidence from Strategy Surveys." *Economic Inquiry.*

Bruine de Bruin, W. and K. G. Carman. 2012. "Measuring Risk Perceptions: What Does the Excessive Use of 50% Mean?" *Medical Decision Making*, 32: 232-236.

Budescu, D. V., and T.S. Wallsten. 1995. "Processing Linguistic Probabilities: General Principles and Empirical Evidence." In J. Busemeyer, R. Hastie, & D. L. Medin (Eds.), *Decision Making from a Cognitive Perspective, Pyschology of Learning and Motivation: Advances in Research and Theory* (Vol. 32, pp. 275–318).

Cai, Hongbin, Yuyu Chen, and Hanming Fang. 2009. "Observational Learning: Evidence from a Randomized Natural Field Experiment." *American Economic Review*, 99(3): 864-82.

Camerer, Colin F, Teck-Hua Ho, and Juin-Kuan Chong. 2004. "A cognitive hierarchy model of games." *Quarterly Journal of Economics*, 861–898.

Camerer, Colin, George Loewenstein, and Drazen Prelec. 2005. "Neuroeconomics: How neuroscience can inform economics." *Journal of Economic Literature* 43, 9–64.

Caminal, Ramon, and Xavier Vives. 1996. "Why market shares matter: An information-based theory." *RAND Journal of Economics*: 221-239.

Caplin, Andrew, and Mark Dean. 2011. "Search, Choice, and Revealed Preference." *Theoretical Economics,* Volume 6, Issue 1, pages 19–48.

Caplin, Andrew, and Mark Dean. 2015. "Revealed Preference, Rational Inattention, and Costly Information Acquisition." *American Economic Review*,, 105(7): 2183-2203.

Caplin, Andrew, Mark Dean, Paul W Glimcher, Robb B Rutledge. 2010. "Measuring beliefs and rewards: a neuroeconomic approach." *Quarterly Journal of Economics.*

Caplin, Andrew, Mark Dean and John Leahy. 2016a. "Consideration Sets and Rational Inattention." NYU Working paper.

Caplin, Andrew, Mark Dean and John Leahy. 2016b. "Rationally Inattentive Behavior." NYU Working paper.

Caplin, Andrew, Mark Dean, and Daniel Martin. 2011. "Search and Satisficing." *American Economic Review*, 101: 2899-2922.

Caplin, Andrew, John Leahy, and Filip Matějka. 2016. "Rational Inattention and Inference from Market Share Data." NYU Working Paper.

Caplin, Andrew, Mi Luo, and Kathleen McGarry. 2016. "Measuring and Modeling Intergenerational Links in Relation to Long-Term Care." NYU Working paper.

Caplin, Andrew, and Daniel Martin. 2015. "A Testable Theory of Imperfect Perception." *Economic Journal* 125.582: 184-202.

Caplin, Andrew, and Daniel Martin. 2016. "The Dual-Process Drift Diffusion Model: Evidence from Response Times." *Economic Inquiry*, 54 (2): 1274-1282.

Carson, Rachel. 2012. "Contingent Valuation: A Practical Alternative when Prices Aren't Available." *Journal of Economic Perspectives,* 26.4, 43-56

Case, Karl E., Robert J. Shiller, and Anne K. Thompson. 2012. "What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets." *Brookings Papers on Economic Activity*, 265–315.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and taxation: Theory and evidence."*American Economic Review* 2009, 99:4, 1145–1177

Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv. 2007. "Consistency and Heterogeneity of Individual Behavior under Uncertainty." *American Economic Review* 97, no. 5: 1921-1938.

Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman. 2014. "Who is (more) rational?" *American Economic Review* 104, no. 6: 1518-1550.

Cover, T.M. and Thomas, J.A., 2012. *Elements of Information Theory.* John Wiley & Sons.

Crawford, Ian. 2010. "Habits Revealed." *Review of Economic Studies*, 77(4), pp. 1382-1402

Dayan, Peter, and Nathaniel D. Daw. 2007. "Decision theory, reinforcement learning, and the brain." *Cognitive, Affective, & Behavioral Neuroscience* 8.4: 429-453.

Daw, Nathaniel D., and Kenji Doya. 2006. "The computational neurobiology of learning and reward." *Current Opinion in Neurobiology* 16.2: 199-204.

Dean, Mark, and Daniel Martin. 2015. "Measuring Rationality with the Minimum Cost of Revealed Preference Violations." *Review of Economics and Statistics.*

Dean. Mark, and Nathaniel Neligh. 2016. "Experminental Tests of Rational Inattention." Columbia University Working Paper.

De Finetti, Bruno. 1965. "Methods for Discriminating Levels of Partial Knowledge Concerning a Test Item." *British Journal of Mathematical and Statistical Psychology* 18.1: 87-123.

De Nardi, Christine. 2004. "Wealth Inequality and Intergenerational Links." *Review of Economic Studies*, 71(3), 743-768.

Delavande, Adeline, and Charles F. Manski. 2010 "Probabilistic Polling And Voting In The 2008 Presidential Election Evidence From The American Life Panel." *Public Opinion Quarterly* (2010): nfq019.

Delavande, Adeline, and Susann Rohwedder. 2008. "Eliciting Subjective Probabilities in Internet Surveys." *Public Opinion Quarterly* 72.5: 866-891.

Dixit, Avinash. 2102. "Paul Samuelson's Legacy." *Annual Review of Economics* Vol. 4: 1-31 (Volume publication date September 2012)

Dominitz, J., and Charles F. Manski. 1996. "Eliciting Student Expectations of the Returns to Schooling." *Journal of Human Resources*, 31, 1–26.

Dominitz, J., and Charles F. Manski. 1997. "Using Expectations Data to Study Subjective Income Expectations." *Journal of the American Statistical Association*, 92, 855–867

Duflo, E. and E.Saez, 2003. "The role of information and social interactions in retirement plan decisions: evidence from a randomized experiment." *Quarterly Journal of Economics* 118(3), 815–842.

Dynan, Karen, Jonathan Skinner, and Stephen Zeldes. 2002. "The Importance of Bequests and Life-Cycle Saving in Capital Accumulation: A New Answer." *American Economic Review Papers and Proceedings*, 274-278.

Enke, Benjamin, and Florian Zimmermann. 2013. "Correlation Neglect in Belief Formation." Mimeo.

Eyster, Erik, and Matthew Rabin. 2010. “Naive Herding in Rich-Information Settings.” *American Economic Journal: Microeconomics*, 2(4): 221–243.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, Uwe Sunde, and David Huffman. 2015. “The Nature of Human Preferences: Global Evidence.” Working Paper.

Falmange, J. C. 1971. “The Generalized Fechner’s Problem and Discrimination.” *Journal of Mathematical Psychology*, 1971, 8, 22-43.

Fehr, Ernst, and Antonio Rangel. 2011. “Neuroeconomic Foundations of Economic Choice - Recent Advances.” *Journal of Economic Perspectives*, 25(4): 3-30.

Finkelstein, A. and K. McGarry, 2006. “Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market.” *American Economic Review*, 96(4): 938-958.

Fischhoff, B. and W. Bruine De Bruin. 1999. “Fifty–fifty= 50%?” *Journal of Behavioral Decision-Making*, 12(2): 149-163.

Fuster, Andreas, David Laibson and Brock Mendel. 2010. "Natural Expectations and Macroeconomic Fluctuations." *Journal of Economic Perspectives*, 24(4):67-84.

Foster, A. and M.Rosenzweig, M. 1995. “Learning by doing and learning from others: human capital and technical change in agriculture.” *Journal of Political Economy* 103(6), 1176–1209.

Frydman, C., N. Barberis, C. Camerer, P. Bossaerts, and A. Rangel. 2014. “Using Neural Data to Test a Theory of Investor Behavior: An Application to Realization Utility.” The Journal of Finance, 69(2), 907-946.

Fudenberg, Drew; Strack, Philipp; Strzalecki, Tomasz. 2015. “Stochastic Choice and Optimal Sequential Sampling.” eprint arXiv:1505.03342

Gabaix, Xavier. 2014. "A Sparsity-Based Model of Bounded Rationality." *Quarterly Journal of Economics*, 129(4): 1661–1710.

Gabaix, Xavier, David Laibson, Guillermo Moloche and Stephen Weinberg, 2006. "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model." A*merican Economic Review*, Vol. 96(4), pages 1043-1068.

Geng, Sen. 2015. "Decision Time, Consideration Time, and Status Quo Bias." *Economic Inquiry.*

Glaeser, Edward L., and Charles G. Nathanson. 2015. "An Extrapolative Model of House Price Dynamics." NBER Working Paper No. 21037/

Glimcher, Paul. 2011. *Foundations of Neuroeconomic Analysis.* New York: Oxford.

Gneezy, Niederle, and Rustichin., 2003, "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics*, CXVIII, 1049 – 1074.

Gul, Faruk, and Wolfgang Pesendorfer. 2006. "Random expected utility." *Econometrica* 74.1: 121-146.

Giustinelli, Pamela. 2012. "Group Decision Making with Uncertain Outcomes: Unpacking Child-Parent Choice of the High School Track." *International Economic Review.*

Gureckis, Todd M., and Douglas B. Markant. 2012. "Self-directed learning a cognitive and computational perspective." *Perspectives on Psychological Science* 7.5: 464-481.

Haavelmo, Trygve. 1958. "The role of the econometrician in the advancement of economic theory." *Econometrica* 351-357.

Heckman, James J., and Bo E. Honore. 1990. "The empirical content of the Roy model." *Econometrica*: 1121-1149.

Holt, Charles A., and Susan K. Laury. 2002. “Risk aversion and incentive effects.” *American Economic Review* 92.5 (2002): 1644-1655.

Hommes, Cars. 2013. *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*. Cambridge University Press

Hossain, Tanjim, and Ryo Okui. 2013. “The binarized scoring rule." *Review of Economic Studies* 80.3: 984-1001.

Hudomiet, P., G. Kézdi and R. J. Willis. 2011. “Stock Market Crash and Expectations of American Households.” *Journal of Applied Econometrics*, 26(3): 393-415. PMC3089779

Hudomiet, P. and R. J. Willis. 2013. “Estimating Second Order Probability Beliefs from Subjective Survival Data.” *Decision Analysis*, 10(2): 152-170. PMC3882032

Hurd, M. 2009. “Subjective Probabilities in Household Surveys.” *Annual Review of Economics*, 1: 543-562.

Hurd, Michael D., and Kathleen McGarry. 1995. “Evaluation of the subjective probabilities of survival in the health and retirement study.” *Journal of Human Resources*: S268-S292.

Hurd, Michael D., and Kathleen McGarry. 2002. “The predictive validity of subjective probabilities of survival.” *Economic Journal* 112.482: 966-985.

Hurd, M. and S. Rohwedder. 2012. “Stock Price Expectations and Stock Trading.” NBER Working Paper 17973.

Juster, F. Thomas. 1966. “Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design.” *Journal of the American Statistical Association* 61.315: 658-696.

Kesternich, Iris; Heiss, Florian; McFadden, Daniel; Winter, Joachim. 2012. “Suit the Action to the Word, the Word to the Action: Hypothetical Choices and Real Decisions in Medicare Part D." Munich Discussion Paper, No. 2012-25"

Kézdi, G. and R. J. Willis. 2011. "Household Stock Market Beliefs and Learning." NBER Working Paper 17614.

Kimball, Miles S., Claudia R. Sahm, and Matthew D. Shapiro. 2008. "Imputing risk tolerance from survey responses." *Journal of the American statistical Association* 103.483: 1028-1038.

Koijen, Ralph , Stijn van Niewerburgh, and Motohiro Yogo. 2015. "Health and mortality delta: Assessing the welfare cost of household insurance choice." *Journal of Finance.*

Krajbich, Ian, Dingchao Lu, Colin Camerer, and Antonio Rangel. 2012. "The Attentional Drift-diffusion Model Extends to Simple Purchasing Decisions." *Frontiers in Psychology*, 3: 193.

Krajbich, Ian, and Antonio Rangel. 2011. "Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions." *Proceedings of the National Academy of Sciences* 108.33: 13852-13857.

Kuhnen, Camelia M. 2015. "Asymmetric Learning from Financial Information." *Journal of Finance* Volume 70, Issue 5, pages 2029–2062.

Lockwood, Lee. 2012. "Bequest Motives and the Annuity Puzzle." *Review of Economic Dynamics*, Vol. 15, Pages 226–243.32

Luce, R. Duncan. 1956. "Semiorders and a Theory of Utility Discrimination." *Econometrica*: 178-191.

Luce, R. Duncan. 1958. "A probabilistic theory of utility." *Econometrica*: 193-224.

Luce, R. Duncan and J. Tukey. 1964. "Simultaneous conjoint measurement: A new type of fundamental measurement." *Journal of Mathematical Psychology*, 1, 1-27.'

Malmendier, Ulrike, and Stefan Nagel. 2016. "Learning from inflation experiences." *Quarterly Journal of Economics.*

Manski, Charles F. 1990. "The Use of Intentions Data to Predict Behavior: A Best-case Analysis." *Journal of the American Statistical Association* 85.412: 934-940.

Manski, Charles F. 1999. "Analysis of Choice Expectations in Incomplete Scenarios." *Journal of Risk and Uncertainty*, 19: 49-66.

Manski, Charles F. 2004. "Measuring Expectations." *Econometrica*: 1329-1376.

Manski, Charles F. and F. Molinari. 2010. "Rounding Probabilistic Expectations in Surveys." *Journal of Business and Economic Statistic*s, 28(2): 219–231. PMC2847785

Manzini, Paola, and Marco Mariotti. 2007. "Sequentially Rationalizable Choice." *American Economic Review*, 97(5): 1824-1840.

Manzini, Paola; Mariotti, Marco. 2014. "Stochastic choice and consideration sets." *Econometrica*, Vol. 89, No. 3, p. 1153-1176.

Marschak, Jacob. 1974. "Limited Role of Entropy in Information Economics." *Theory and Decision* 5.1: 1-7.

Martin, Daniel. 2016. ""Consumer Theory with Price Uncertainty." Working Paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2393035

Masatlioglu, Yusufcan, Daisuke Nakajima and Erkut Ozbay. 2011. "Revealed Attention." *American Economic Review*, 102(5): 2183-2205.

Matějka, Filip, and Alisdair McKay. 2015. "Rational Inattention to Discrete Choices: a New Foundation for the Multinomial Logit Model." *American Economic Review*, vol. 105 (1), p. 272-298.

McFadden, D. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." in P. Zarembka, ed, *Frontiers of Econometrics*, Academic Press: New York, 105-142.

McFadden, D. 1986. "The Choice Theory Approach to Market Research." *Marketing Science*, 5, 275-297.

Munshi, K. 2004. "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution." *Journal of Development Economics* 73(1), 185–213.

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *American Economic Review*: 1313-1326.

Niederle, Muriel and Lise Vesterlund. 2010. "Explaining the Gender Gap in Math Test Scores: The Role of Competition," *Journal of Economic Perspectives*. Vol. 24(2), 129–144.

Nyarko, Yaw, and Andrew Schotter. 2002. "An Experimental Study of Belief Learning using Elicited Beliefs." *Econometrica* 70.3 971-1005.

Odean, Terrance. 1998. "Are investors reluctant to realize their losses?" *Journal of Finance* 53.5: 1775-1798.

Ratcliff, Roger. 1978. "A Theory of Memory Retrieval." *Psychological Review*, 85(2): 59.

Ratcliff, Roger, Philip L. Smith, Scott D. Brown, Gail McKoon. 2016. "Diffusion Decision Model: Current Issues and History Review Article." *Trends in Cognitive Sciences*, http://www.sciencedirect.com/science/article/pii/S1364661316000255

Rietveld C. A.,. . . , P. M. Visscher, D. J. Benjamin, D. Cesarini & P. D. Koellinger. 2014. "Common Genetic Variants Associated with Cognitive Performance Identified Using Proxy-Phenotype Method." *Proceedings of the National Academy of Sciences* 111: 13790-13794.

Rubinstein, Ariel. 2007. “Instinctive and Cognitive Reasoning: A Study of Response Times.” *Economic Journal*, 117: 1243-1259.

Salant, Yuval, and Ariel Rubinstein. 2008. “(A, f): Choice with Frames.” *Review of Economic Studies* 75.4, 2008. 1287-1296.

Samuelson, Paul A. 1938. “A Note on the Pure Theory of Consumer’s Choice.” *Economica*, 5:61-71.

Savage, Leonard J. 1971. “Elicitation of personal probabilities and expectations.” *Journal of the American Statistical Association* 66.336 (1971): 783-801.

Schotter, Andrew, and Isabel Trevino. 2014. “Belief elicitation in the laboratory.” *Annual Review of Economics* 6.1: 103-128.

Schultz, Wolfram, Peter Dayan, and P. Read Montague. 1997. “A neural substrate of prediction and reward.” *Science* 275.5306: 1593-1599.

Shadlen, Michael N., and William T. Newsome. 2001. “Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey.” *Journal of Neurophysiology* 86.4: 1916-1936.

Simon, Herbert A. 1955. “A behavioral model of rational choice.” *Quarterly Journal of Economics*: 99-118.

Sims, Christopher A. 1998. “Stickiness.” *Carnegie-Rochester Conference Series on Public Policy* 49, 317–356.

Sims, Christopher A. 2003. “Implications of Rational Inattention.” *Journal of Monetary Economics* 50, 665–690.

Society of Actuaries. 2014. Land This Plane: A Delphi Research Study of Long-Term Care Financing Solutions.

Sorensen, A. 2006. Social learning and health plan choice.” RAND Journal of Economics 37(4), 929–945.

Spiliopoulos, Leonidas, and Andreas Ortmann. 2014. "The BCD of Response Time Analysis in Experimental Economics." Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2401325.

Stigler, George J. 1961. "The Economics of Information." *Journal of Political Economy*: 213-225.

Tetlock, P. E. and B. A. Mellers. 2011. "Intelligent Management of Intelligence Agencies: Beyond Accountability Ping-Pong." *American Psychologist*, 66(6): 542-554. PMID: 21823782

Thurstone, L. L. 1931. "The indifference function." *Journal of Social Psychology*, 2, 139-167.

Thurstone, L. L. 1927, "A Law of Comparative Judgement." *Psychological Review*, 34, 273-286.

Van der Klaauw, Wilbert. 2012. "On the use of expectations data in estimating structural dynamic choice models." *Journal of Labor Economics*, 30(3), 521-554.

Varian, Hal R. 1982. "The Nonparametric Approach to Demand Analysis." *Econometrica*: 945-973.

Weber, E. H. 1834. *On the Tactile Senses*. H.E. Ross and D.J.Murray Trans. and Ed. New York: Experimental Psychology Series.

Wilcox, Nathaniel T. 1993. "Lottery Choice: Incentives, Complexity and Decision Time." *Economic Journal*: 1397-1417.

Wiswall, Matthew, and Basit Zafar. 2015a. "Determinants of College Major Choice: Identification using an Information Experiment." *Review of Economic Studies,* 82.2: 791-824.

Wiswall, Matthew, and Basit Zafar. 2015b. "How Do College Students Respond to Public Information about Earnings?" *Journal of Human Capital* 9.2: 117-169.

Wiswall, Matthew, and Basit Zafar. 2016. "Preference for the workplace, investment in human capital, and gender."

Woodford, Michael. 2009. "Information-Constrained State-Dependent Pricing," *Journal of Monetary Economics* 56, 100-124.

Woodford, Michael. 2014. "An Optimizing Neuroeconomic Model of Discrete Choice." *National Bureau of Economic Research*, Cambridge, Mass. Working Paper 19897.