How Are Preferences Revealed?

John Beshears  
Harvard University

James J. Choi  
Yale University and NBER

David Laibson  
Harvard University and NBER

Brigitte C. Madrian  
Harvard University and NBER


Current draft: January 15, 2007

The authors are grateful to participants at the “Happiness and Public Economics Conference” for their feedback, particularly Alan Manning.
How Are Preferences Revealed?

Abstract: It sometimes makes sense to assume that revealed preferences (preferences that rationalize an economic agent’s observed actions) and normative preferences (preferences that represent an economic agent’s actual interests) are identical. But there are many cases where this assumption is violated. We first identify five factors that increase the likelihood of a disparity between revealed preferences and normative preferences: passive choice, complexity, limited personal experience, third-party marketing, and intertemporal choice. We then discuss six approaches that jointly contribute to the identification of normative preferences: structural estimation, active decisions, asymptotic choice, aggregated revealed preferences, reported preferences, and informed preferences. Each of these approaches relies on consumer behavior to infer some property of normative preferences without equating revealed and normative preferences. We illustrate these issues with evidence from savings and investment outcomes.

John Beshears  
Department of Economics  
Harvard University  
Littauer Center  
Cambridge, MA 02138  
beshears@fas.harvard.edu

James J. Choi  
Yale School of Management  
135 Prospect Street  
P.O. Box 208200  
New Haven, CT 06520-8200  
james.choi@yale.edu

David Laibson  
Department of Economics  
Harvard University  
Littauer Center  
Cambridge, MA 02138  
dlaibson@harvard.edu

Brigitte C. Madrian  
John F. Kennedy School of Government  
Harvard University  
79 JFK Street  
Cambridge, MA 02138  
Brigitte_Madrian@harvard.edu
Revealed preference theory shows how to construct utility functions from price and choice observations (Samuelson, 1938, 1948; Little, 1949; Houthakker, 1950; Afriat, 1967). Economists nearly always assume that these revealed preferences are normative preferences—preferences that represent the economic actor’s true interests. And when economists make policy recommendations, they are typically made based upon revealed preferences, implicitly assuming that revealed preferences and normative preferences are identical.

In some situations, it makes sense to give revealed preferences normative status. When a ten-year-old child chooses chocolate over vanilla ice cream, she reveals a legitimate preference for one flavor over another. The government should respect that preference, and ice cream parlors should cater to it.

However, there are many cases in which even the choices of adult consumers do not reveal a true preference, but rather reflect the combined influence of true preferences and decision-making errors. When a worker invests all of her retirement savings in her employer’s stock, economists should not assume that she has acted in her best interest. In surveys, workers on average report that an investment in their employer’s stock is less risky than an investment in a diversified mutual fund (John Hancock Financial Services, 2002). When workers put all their financial eggs in one basket, they are probably revealing many things, one of which is confusion about the true risk characteristics of employer stock.

Human behavior is jointly determined by both normative preferences and other factors such as analytic errors, myopic impulses, inattention, passivity, and misinformation, to name a few. Despite these complications, economists need not throw up their hands and reject all revealed preferences. This paper describes a compromise in which economists use behavior to identify normative preferences but take care to acknowledge the occasional wedge between revealed preferences and normative preferences. Naturally, we are not alone in worrying about this distinction (e.g. Luce, 1959; Bernheim and Rangel, 2006).

In Section 1, we describe the kinds of situations where normative preferences are least likely to be revealed by people’s choices. We illustrate these issues with evidence from savings
and investment decisions. We identify five factors that increase the likelihood that revealed preferences will not have normative merit: passive choice, complexity, limited personal experience, third-party marketing, and intertemporal choice.

Section 2 discusses six frameworks that can jointly contribute to the identification of normative preferences: structural estimation, active decisions, asymptotic choice, aggregated revealed preferences, reported preferences, and informed preferences. Each framework relies on consumer behavior to infer some property of normative preferences without equating revealed and normative preferences. Practical policy analysis should use insights from all of these frameworks.

1. Red flags

In this section we discuss five factors—passive choice, complexity, limited personal experience, third-party marketing, and intertemporal choice—that often create a wedge between revealed and normative preferences.

1.1. Passive choice

In many situations, economic agents do not actively make choices. Instead, they passively accept defaults that are chosen by others. In theory, acceptance of a default could be a conscious and meaningful decision. In practice, acceptance of a default often reflects other forces, like procrastination or the power of suggestion. Many agents who procrastinate will temporarily accept defaults because they plan to opt out of the default at some later date. However, it can take years to implement these subsequent actions, even when the economic consequences of delay are non-trivial (Choi, Laibson and Madrian, 2005b). Highly suggestible agents may also accept defaults because they believe—perhaps erroneously—that the person who set the default was making a carefully considered recommendation (Madrian and Shea, 2001; Beshears et al., 2006a).

Preferences revealed through passive choice are often unstable, since variation in defaults can generate tremendous variation in outcomes. For example, Choi, Laibson, Madrian and Metrick (2006) examine four companies that changed the default enrollment regime for their

---

employer-sponsored savings plan. When the default was non-participation (an opt-in default), enrollment rates ranged from 25% to 43% six months after hire (even though non-participating employees were foregoing large employer matching contributions at three of the four firms in addition to tax benefits). By comparison, when the default was changed to automatic enrollment in the savings plan unless employees opted-out, enrollment rates six months after hire increased substantially, to 86% to 96%. The within-firm participation increase due to the change in the participation default ranged from 50 to 67 percentage points across the four firms.

The impact of defaults on savings outcomes extends well beyond savings plan participation. Similar effects have been documented for almost every aspect of retirement saving, including asset allocation outcomes, contribution rates, pre-retirement cash distributions, and annuitization (see Beshears et al., 2006a, for a summary of this research).

1.2. Complexity

Complex problems are, by definition, hard to solve. Difficult decisions tend to delay choice, increasing the fraction of individuals that accept default options (O’Donoghue and Rabin, 1999a). Complexity also biases choice, since people tend to avoid more complicated alternatives (Shafir et al., 1993; Iyengar and Kamenica, 2006). Finally, complexity adds noise to choices. Since complicated options may not be well understood, some consumers choosing a complicated option will do so because they misestimate its value (Luce, 1959; McFadden, 1981; Gabaix, Laibson, and Li, 2005).

1.2.1. Complexity and the number of choices

In the savings domain, one factor contributing to complexity is the number of asset allocation options that are available. At year-end 2005, there were 8,454 mutual funds registered in the United States. Allocating funds across these investment options is a difficult task unless the agent has a heuristic that reduces the complexity such as only considering the funds at one trusted mutual fund company. Iyengar and Kamenica (2006) show that experimental subjects are more likely to choose a simple investment rather than a complex investment as the number of investment menu options increases. Iyengar and Kamenica also find support for this result in a

---

cross-section of 401(k) plans: for every 10 additional mutual funds in a 401(k) investment menu, the allocation to equities is 3.3 percentage points lower.

A superabundance of options also delays savings plan enrollment. A typical employer-sponsored savings plan allows employees to contribute up to 15 or 20 percent of pay (subject to IRS contribution limits) in every convex combination of, on average, 14 different investment options (U.S. Bureau of Labor Statistics, 2005; Hewitt Associates, 2005). Moreover, when employees must opt into savings plan participation, decisions about how much to contribute and how to allocate these contributions must be made in order to enroll. Iyengar, Huberman and Jiang (2004) document a negative relationship between the number of investment options offered in a savings plan and employee participation: each additional 10 funds in the menu of investment options is associated with a decline in participation rates of 1.5 to 2.0 percentage points. Conversely, Choi, Laibson, and Madrian (2006a) and Beshears et al. (2006b) show that simplifying the savings plan enrollment decision increases participation rates. They study an intervention at two different firms in which non-participating employees were given the opportunity to opt into a single pre-selected contribution rate and asset allocation. Even though the number of 401(k) choices available did not decrease, this simplified reframing of the enrollment decision increased participation by 10 to 22 percentage points.

1.2.2. Complexity and long horizons

Another factor contributing to the complexity in the savings domain is the time horizon of the task: choices with distant consequences are complex. Simulating what will happen later today is much easier than simulating events on a particular afternoon four decades from now. The challenge of simulating the distant future is closely related to the challenge of simulating any stochastic event. Looking ahead forty years, one needs to consider numerous divergent paths: unemployment, sickness, divorce, bankruptcy, bequests, etc.

Decisions with consequences in the distant future reflect both normative preferences and forecasting errors. Some of those forecasting errors will be predictable. For example, recently released estimates from the Centers for Medicare and Medicaid Services project a 7.2% annual increase in health care spending for the years 2005-2015, outstripping both the annual rates of GDP growth and overall inflation (Borger et al., 2006). It is likely that many households are not aware of these forecasts. The inferences that economists draw regarding households’
intertemporal preferences depend critically on whether households foresee these health costs (some of which they will need to bear). If economists observe a low current savings rate, that could reflect either a high discount rate and a rational forecast of future health costs or a low discount rate and an under-estimate of future health costs. It is hard for the social scientist to know which mechanism is at play.

1.3. Limited personal experience

Limited personal experience creates another wedge between revealed preferences and normative preferences. Human learning is often generated by feedback. A child learns that hot food burns the roof of his mouth through experience rather than lectures. Likewise, credit card account holders learn to pay their bills on time by first paying late fees (Agarwal et al., 2006).

Consumers with little or no feedback are not likely to learn what is in their best interest. What personal experiences could teach a middle-aged worker whether she is saving the right amount for retirement? In principle, she could learn by observing others, particularly people in other generations. However, every generation faces different financial circumstances. Moreover, people are generally far more responsive to their own experiences than the experiences of others. For example, Choi, Laibson and Madrian (2005a) show that the bankruptcies and ensuing 401(k) debacles at Enron, WorldCom and Global Crossing did little to reduce the amount invested in employer stock in the savings plans of workers at other U.S. firms.

1.4. Third-party marketing

Some behavior is influenced by marketing. Tom Sawyer tricked his friends into paying him for the privilege of painting his family’s fence. Such preference distortions do not arise exclusively in the world of fiction. For example, S&P 500 index funds charge fees that vary by an order of magnitude (Hortaçsu and Syverson, 2004). Moreover, this range of fees cannot be explained by variation in bundled non-portfolio services (Choi, Laibson, and Madrian, 2006b). Instead, the willingness to pay high fees is partially due to the effective marketing and branding of the underlying good—an index fund—which is a commodity.

It is not obvious how economists should evaluate preferences for a branded commodity. There are some cases where such preferences may not deserve normative weight. When asset management firms induce their clients to invest in dominated (high-fee) assets and when
employers persuade their rank-and-file workers to hold employer stock, economists should wonder whether these revealed preferences have normative legitimacy or reflect Tom Sawyer effects (Ariely et al., 2003).

1.5. Intertemporal choice

Intertemporal choices raise two sets of normative questions. First, even if an economic agent makes choices that imply a consistent discount rate, it is not obvious that economists should give that revealed discount rate normative weight (Ramsey, 1928; Pigou, 1932; Harrod, 1948; Solow, 1974; Kahneman et al., 1997; Anand and Sen, 2000). Why should utils experienced at date $t$ have less weight than utils experienced at a later date $t+\tau$? Only discounting due to mortality risk seems to be easily defended philosophically. However, mortality discounting is orders of magnitude too small to explain the intertemporal preferences revealed by actual households’ savings choices; the revealed exponential discount rate is around 5% per year (e.g., Gourinchas and Parker, 2002). Recently, Nordhaus (2006) has shown that the policy recommendations for addressing global warming hinge critically on the question of the normative discount rate. For example, the highly influential Stern report (2006) bases its analysis on a normative discount rate of 0.1% per year, which is approximately $1/50^{th}$ the magnitude of the revealed preference discount rate of 5%.

3 The discussion in the previous paragraph posits the existence of a constant revealed discount rate, and then asks whether the normative discount rate should match it. However, a growing body of evidence suggests that revealed discount rates are not constant. Instead, discount rates are higher in the short run than the long run (Ainslie, 1992; Loewenstein and Prelec, 1992; Laibson, 1997; Angeletos et al., 2001; DellaVigna and Malmendier, 2006). Such variation raises a host of additional normative problems. One particularly salient problem is that non-constant discount rates imply dynamically inconsistent preferences.

To illustrate this problem, imagine that a person can make an investment at a cost of $C$ utils to gain delayed benefits of $B$ utils. For analytic simplicity, assume that benefits occur one

---

3 From the FAQ section of the Stern report: “We carefully examine the case for discounting the future just because it is the future—which in economic terms is known as pure time preference. This requires a consideration of the ethical issues involved in comparing the incidence of costs and benefits between generations, some of which are distant in time. We argue—in line with economists including Ramsey, Sen, Pigou and Solow—that the welfare of future generations should be treated on a par with our own. This means that the only justification for a positive rate of pure time preference in assessing the impacts of climate change is the possibility that the human race may be extinguished. As the possibility of this happening is low, we assume a low rate of pure time preference, 0.1%.”
period after the costly investment. The individual has a quasi-hyperbolic discounting function (Phelps and Pollak, 1968; Laibson, 1997), whereby rewards and costs at times 0, 1, 2, 3, ..., are multiplied by respective weights $1, \beta \delta, \beta^2 \delta^2, \beta^3 \delta^3, \ldots$, with $0 < \beta \leq 1$ and $0 < \delta \leq 1$. When $\beta = 1$, this model is identical to the classical exponential discounting model. When $\beta < 1$, this model reproduces the “hyperbolic” pattern of more discounting in the short run and less discounting in the long run. O’Donoghue and Rabin (1999b) call these preferences “present-biased.”

Consider the case in which $\beta = 1/2$ and $\delta = 1$ (the value $\delta = 1$ is used in Akerlof, 1991). Set $C = 4$ and $B = 6$, so the undiscounted cost of investment is 4 and the undiscounted benefit is 6. With these parameter values, the investment looks appealing when viewed from a temporal distance:

$$-\beta \delta^t C + \beta \delta^{t+1} B = -(1/2)(4) + (1/2)(6) = 1 > 0.$$ 

However, the investment is undesirable if the agent is asked to do it right now:

$$-C + \beta \delta B = -(4) + (1/2)(6) = -1 < 0.$$ 

Hence, a person with this discounting function has two competing sets of revealed preferences. When asked to make binding commitments in advance, the agent will choose to invest. When such binding commitments are not available, the agent will not end up investing (since she always breaks her previous plans when the moment of action arises). In situations like this, revealed preferences cannot be a reliable guide to normative preferences. Alternative or additional normative assumptions are needed.

2. Inferring normative preferences

The previous section described situations in which revealed preferences deviate from normative preferences. For policy purposes, economists would like to be able to identify normative preferences when such deviations occur. In the current section, we summarize six methods for doing so. None of these methods is perfect, and all of them should be viewed as complementary ways of measuring normative preferences. As we discuss each method, we will review its strengths and weaknesses.
2.1. **Structural estimation**

For an economist trained in modern empirical methods, a natural way to measure normative preferences is to estimate them with a structural model. Such a model would have two components:

1. A (positive) behavioral model with a parameter vector \( \theta \)
2. A set of normative axioms that map the parameters \( \theta \) into normative preferences

The behavioral model would incorporate all of the economic and psychological motives that shape behavior, including the mechanisms that generate mistakes (e.g., non-Bayesian inference) and the preferences that are normatively illegitimate (e.g., present bias\(^4\)).

To implement this analysis, experimental data or field data would be used to estimate the behavioral model (1) and the associated parameter vector \( \theta \). These parameters would then be mapped into normative preferences using the axioms in (2).

As an example, suppose consumers’ static revealed preferences are normatively legitimate but their dynamic preferences are not, as in the case where they suffer from present bias. Such an argument would be supported by the normative axiom that the welfare function should have no intertemporal discounting except for that due to mortality risk (Ramsey, 1928; Pigou, 1932; Harrod, 1948; Solow, 1974; Kahneman et al., 1997; Anand and Sen, 2000). At the same time, the normative axioms might respect people’s static preferences over risky alternatives. In this example, \( \theta \) would contain both static risk aversion parameters and dynamic discounting parameters. Both types of parameters would be simultaneously estimated from available data,\(^5\) but only the static risk preferences would be elevated to normative status.

This structural estimation framework has the strength that it forces the researcher to make clear assumptions about behavioral biases while enabling the researcher to simultaneously identify both normative preferences and the underlying behavioral model. On the other hand, the structural model has the weakness that it is costly to implement—like most structural models—and it may not be robust to model specification errors.

---

\(^5\) See Laibson, Repetto, and Tobacman (2006) for an example of such a structural estimation, absent any normative interpretation.
2.2. Active decisions

In some cases, revealed preferences and normative preferences diverge because people stay at a default that they report is not in their best interest. For example, two-thirds of survey respondents at one company reported that their current savings rate was too low relative to their ideal savings rate. A third of these undersavers said they were planning to increase their savings plan contribution rate in the next two months, but almost none of them actually did so (Choi, Laibson, Madrian, and Metrick, 2002 and 2006).

For such workers, the preference they reveal through their action is that saving more is not (yet) optimal: they do not increase their savings plan contribution rate. However, their stated preference contradicts this revealed preference. For such conflicted households, it is not clear what their normative preferences actually are. Since their self-reports deviate from their revealed preferences, economists should be hesitant to accept their revealed preferences at face value.

In situations like this, it is illuminating to examine how behavior varies with the type of default that is used. As discussed above, savings plan participation rates vary enormously with the participation default that is chosen: opt-in versus opt-out. The defaults biases outcomes, both because consumers tend to procrastinate and because consumers tend to follow the implicit advice that a default represents. A default-free mechanism might be less biased. For example, an active decision mechanism forces individuals to explicitly state their own preferences (Choi, Laibson, Madrian, and Metrick, 2005). In the context of an employer sponsored savings plan, such an approach might require all workers to affirmatively state whether they do or do not want to participate in the savings plan by a certain deadline. Choi et al. document that the enrollment rates achieved under an active decision regime tend to be between those obtained under either of the arguably biased default regimes (although closer to the participation rates achieved under automatic enrollment).

2.3. Asymptotic choice

In most stationary economic environments, short-run choices are likely to be further from normative optimality than long-run choices. For example, in firms with saving plans that use a non-enrollment (opt-in) default, procrastination may delay enrollment, thereby biasing down the short-run participation rate. Learning mechanisms—including imitation of peers, formal
education, and experiential feedback—may also advantage long-run choices relative to short-run choices.

Hence, it is likely that the participation/savings behavior of experienced workers is normatively superior to the behavior of inexperienced workers. One should therefore give disproportionate weight to the investment behavior of highly experienced—i.e., high-tenure—workers when attempting to infer normative preferences (Choi, Laibson, Madrian, and Metrick, 2003 and 2005).

Credit card data contain direct and detailed evidence of learning effects. Controlling for person fixed effects and transitory characteristics, new account holders pay fees that are five times greater than the fees paid by experienced account holders (Agarwal et al., 2006). The fees paid by new account holders reflect mistakes that are avoided by more experienced account holders. Each time an account holder pays a fee—for instance, a late payment fee—her likelihood of paying that fee the next month drops by a third.

When new account holders pay fees, most are not revealing a high willingness to pay fees in exchange for services (like paying $30 for the service of being able to pay one’s bill late). Instead, new account holders are revealing that they have not yet optimized their account management skills. With time and experience, such optimization takes place, and fee payments largely vanish. It is this asymptotic behavior that reflects the true preferences of the account holders.

2.4. Aggregated revealed preferences

When homogeneous individuals make noisy, error-prone decisions, their individual decisions will not reflect normative preferences, but their aggregate behavior may. Hence, normative preferences can sometimes be inferred from the central tendencies of aggregate distributions of behavior.

However, aggregate analysis poses problems when the observed heterogeneity in revealed preferences is due both to noisy errors and to heterogeneity in normative preferences. In this case, normative preferences can only be imputed if the researcher specifies a model of heterogeneous preferences and decision-making errors.

---

6 For this analysis, experienced account holders have at least four years of feedback.
Furthermore, if decision errors do not have a zero mean, inferences from a population’s central moments can lead to biased conclusions. The common phenomenon of employer stock holding in employer sponsored savings plans is likely to be a manifestation of a non-zero-mean error.

2.5. Reported preferences

In many cases, decision-makers report that they know what they should do but nevertheless fail to implement that action. For example, 70% of smokers in the U.S. report that they want to quit, and 41% stopped smoking for a day or more during the previous twelve months in an unsuccessful attempt to quit (Centers for Disease Control and Prevention, 2002).

Historically, economists have rejected all self-reports on the grounds that behavior has real consequences and self-reports are (usually) only cheap talk. We agree that self-reports can’t be taken at face value, but we also believe that they should not be ignored completely. In our view, successful models of human decision-making should be able to explain both behavior and self-reports. Like behavior, self-reports can be measured and modeled. Like behavior, self-reports can be used to predict things that economists unambiguously care about (like future behavior).

We recognize that self-reports are often motivated by signaling, but behavior can also be driven by signaling motives. Self-reports reveal at least something about an agent’s goals and values. Normative economics should allow self-reports to have some standing. It would be strange to try to infer someone’s normative preferences without at least considering their own views on the question.

Finally, self-reports can be used to assess a consumer’s confidence that his behavioral choices are optimal. Choi, Laibson, and Madrian (2006b) find that consumers who choose low-fee index funds tend to be more confident about their investment choice than consumers who choose high-fee index funds. Hence, self-reports may provide a natural tool for escaping the tautology of equating revealed preferences and normative preferences. In this view, economists should be more willing to interpret revealed preferences as normative preferences when consumers report that they are confident in their own choices.
2.6. Informed preferences

Informed opinions come in two forms. First, external observers may offer expert advice (e.g., academics, consultants, financial planners, etc.). Second, decision-makers may themselves gain more expertise when they receive formal training and education. When trained/educated decision-makers make a choice, we call this an “informed preference.”

In our view, both sources of expert opinion—external experts and practicing experts—should play an important role in the identification of normative preferences. This is particularly true when exogenous education or training can be shown to systematically change people’s behavior. Naturally, there are limits to such arguments, especially when “education” is little more than marketing or brainwashing. However, at least some education is not overwhelmed by these problems. When economists measure normative preferences, we should give disproportionate weight to the actors who have good reason to know what they are doing.

3. Conclusion

Once economists acknowledge a gap between revealed preferences and normative preferences, it is natural to extend our models so that we can clearly distinguish between them. We have reviewed five factors that tend to increase the gap between revealed and normative preferences: passive choice, complexity, limited personal experience, third-party marketing, and intertemporal choice.

We have also discussed six ways of measuring normative preferences when they deviate from revealed preferences. Structural estimation specifies a positive model with a precise set of economic and psychological motives (perhaps including non-Bayesian thinking and other decision-making errors). This model is then estimated using data, and the resulting positive preferences are mapped into normative preferences using normative axioms.

Active decisions eliminate some biases generated by default regimes. Under an active decision regime, individuals are required to explicitly state their preference without being influenced by a background default. In some circumstances, this preference elicitation will be more reliable (and more socially efficient) than allowing consumers to express their preferences by opting into or out of a default.
In most stationary economic environments, initial choices are likely to be further from normative optimality than choices made after many periods of experience. One should therefore give more weight to asymptotic choices when attempting to infer normative preferences.

When homogeneous individuals make noisy, error-prone decisions, their individual decisions do not reflect normative preferences, but their aggregate behavior can. Hence, normative preferences can sometimes be inferred from the central tendencies of aggregated preferences.

Self-reported preferences reveal something about an agent’s goals and values. Normative economics should allow self-reports to have some standing. This is particularly true when self-reports can be used to distinguish confident consumer decisions from decisions that were made in a state of confusion.

Informed opinions come in two forms. External observers may offer expert advice, and decision-makers may themselves gain more expertise when they receive training or education. When trained/educated decision-makers make a choice, we call this an informed preference. When economists measure normative preferences, we should give disproportionate weight to the actors who are most likely to know what they are doing.
References


