Toward an Agenda for Behavioral Public Finance

Edward J. McCaffery and Joel Slemrod

USC CLEO Research Paper No. C04-22
USC Law and Economics Research Paper No. 04-25

CLEO RESEARCH PAPER SERIES
LAW & ECONOMICS RESEARCH PAPER SERIES

Sponsored by the John M. Olin Foundation

University of Southern California Law School
Los Angeles, CA 90089-0071

This paper can be downloaded without charge from the Social Science Research Network electronic library at http://ssrn.com/abstract=590201
I. Introduction: Notes on a possible field

This essay is about the intersection—or the possible intersection—between the fields of behavioral economics and public finance, which we call behavioral public finance.

Public finance is of course a venerable field in the economics fold. It rests on two basic assumptions or principles. One is that individuals are rational, maximizing agents, in the simple sense that they act consistently on the basis of a single well-defined utility function.1 We call this the “rationality assumption.” Two is that the basis for evaluating social policies should be the well-being of the society’s members, as they judge their well-being to be. We call this the “consumer sovereignty principle.” Behavioral economics, in contrast, is a newcomer on the social science field. It rests on a series of empirical challenges to the rationality assumption that can, if taken to a certain limit, call the consumer sovereignty principle into question. Perhaps because of this perceived nihilistic possibility, or simply because it represents a challenge to received orthodoxy, skeptics continue to question both the facts and the relevance of behavioral economics to any field of study.

---

1 See Becker 62, at 1: “[N]ow everyone more or less agrees that rational behavior simply implies consistent maximization of a well-ordered function, such as a utility or profit function.”
We, however—and however tentatively—believe that to the extent behavioral economics rests on empirically verifiable (and verified) understandings about how real people think, choose, decide, and act in real-life settings, public finance models that aim for real-world relevance ought to take behavioral insights into account. This does not mean a wholesale abrogation of traditional public finance, or an abandonment of consumer sovereignty principles. As in all marriages, there will be give and take; the whole will be different from—and at least potentially better than—the sum of the parts.

After some additional background comments on the two disciplines, we illustrate the possible relevance of behavioral public finance with three broad clusters of questions, respectively concerning the forms of public finance mechanisms, problems of inter-temporal choice, and models of taxpayer compliance.

A. Three faces of public finance

Public finance has been a principal concern of economics at least since Adam Smith. The subject matter of public finance is the role of the government in addressing society’s economic tasks, regarding both the allocation of scarce resources and their distribution or redistribution. As classically divided by Musgrave (1959), public finance consists of two parts, a government expenditure and revenue-raising aspect or, in more colloquial terms, a tax and a spending dimension. Public finance also divides naturally into descriptive or positive analysis and prescriptive or normative analysis. Positive analysis seeks to understand the effect of government tax and expenditure programs on the behavior and well-being of individuals (as well as on economic aggregates such as households or firms). Normative analysis seeks to establish guidelines for what the government should—and should not—do. Answers to fundamental normative questions depend on values to which economic analysis cannot contribute directly, but these answers also depend on the positive analysis. Public finance can lay out the costs and benefits—the welfare implications—of various alternative government actions.

On the spending side, within the neoclassical welfare framework familiar to economists, free, private markets generally work to achieve an efficient outcome. The case for a
particular government expenditure program turns on a demonstration of a market failure, such as the presence of a public good, externality or, increasingly, an informational asymmetry. To these potential efficiency rationales for market intervention, public finance has long added a distributive/redistributive component, whether the income distribution is taken to be a public good\(^2\) or not, because the free market alone does not necessarily produce an attractive distribution of society’s collective resources. In the neoclassical framework public finance seeks both to describe what policies maximize social welfare by alleviating market failures (as by providing public goods) in an allocatively efficient manner, and to redistribute the (greater) social wealth to achieve a more desirable distribution of resources.\(^3\)

The other shoe falls because government expenditure and redistributive programs require government revenues, and so public finance has also long considered the appropriate structure of taxation. Given the infeasibility of first-best or non-distortionary lump-sum taxation, public finance economists have developed sophisticated tools of optimal taxation.\(^4\) But the adoption of optimal tax systems in real-world settings—even bracketing the difficult social choice problems of agreeing on the appropriate welfare function in the first place—is problematic.\(^5\) Public finance economists consider matters of compliance and administration: minimizing the transaction costs of tax. As part of its core mission, public finance studies the behavioral response of individuals and firms to alternative tax systems, seeking to understand and predict both revenue effects and the deadweight costs of various alternative tax regimes and reforms—seeking, that is, to understand the inputs needed for theory to lead to best practices (for example, labor supply elasticities needed to implement optimum income tax policies), and the tradeoffs entailed in second-best, real-world settings.

In addition to these two traditional functions of public finance—tax and spending—government now often pursues a third function through its legally-sanctioned control

---

\(^2\) Thurow 1975.
\(^3\) Mirrlees 1971, Kaplow and Shavell 2002.
\(^5\) Slemrod 1990.
over fiscal matters: it acts to modify behaviors. Under what, if any, circumstances government should engage in this sort of activity is a part of the challenge that behavioral economics poses to traditional economics. Within the neoclassical framework, taxes have long been advocated as a tool to correct for externalities, or (equivalently) the problem of social cost diverging from private cost. But today the government often acts through fiscal mechanisms to encourage what are arguably “goods” (charitable giving, savings) or discourage what are arguably “bads” (smoking, drinking), with or without traditional externality-correcting grounds to guide it. This expansion of the domain of public finance raises questions under and about the consumer sovereignty principle.

B. The New Kid on the Block

In contrast to welfare economics, behavioral economics is a fledgling on the field of social science. It has roots in the seminal work of Herbert Simon (1955) on “bounded rationality,” and grew enormously under the guidance of Daniel Kahneman and Amos Tversky (1979), who argued that there are two broad features of human judgment and decision-making: various errors in coding mechanisms, known as heuristics and biases, that lead to violations of the laws of logic and consistency; and evaluatory functions that differ from the expected utility function of the neoclassical rational choice models. Other treatments (e.g. Thaler 1980) differ in terminology and detail, but all share the characteristic of showing that real people do not follow logically consistent choice and decision protocols. By now abundant experimental and real-world observed evidence—buttressed by common sense—confirms that individuals do not always think and act in ways consistent with the standard, and limited, axioms of rational choice.

With the recent Nobel Prize awarded to Kahneman, the field is blossoming. But it would be premature to say that behavioral economics has yet reached full flower. In particular, the field self-consciously lacks a general “field theory” of human behavior, let

---

6 Pigou 1951. This use of public finance mechanisms was harshly criticized by Coase 1960, 1988.
7 Koszegi and Gruber 2003; O’Donoghue and Rabin 2003; Sunstein and Thaler 2003.
8 von Neumann and Morgenstern 1944, Savage 1954.
9 Camerer 2000.
10 List 2004.
alone one that would be as parsimonious as the rational agent model. Indeed, the field of behavioral economists rests, first and foremost, on a rejection—of the axiom of rational agency—rather than on the affirmation or acceptance of any specific theory of human behavior and thought. Given this, care must be taken in extending behavioral models to new fields. But there are also compelling reasons to go forward.

Consider, for example, the role of various possible structures, institutions, or other mechanisms to vitiate or mitigate the effects of individual heuristics and biases. One set of possibilities includes debiasing mechanisms or techniques—procedures to lay bare illusions and prompt subjects to give consistent responses, including simply revealing inconsistencies and explaining the heuristics and biases commonly in play. A more promising avenue, however, has proven to lie in various “arbitrage mechanisms,” structural devices to counter-balance the effects of irrational biases. The paradigms in private finance are the market and competition itself. Even if some—most?—agents are irrational, and tend to buy high and sell low, markets can be expected to counteract the bias and appropriately value securities, as long as there is at least one rational actor without liquidity constraints. Individual investors may still incur losses but the price system will be on net efficient, diminishing the aggregate harm to social welfare. Similarly, competition in consumer markets keeps prices at marginal cost, however much some individual agents might be able to be tricked into paying more.

Arbitrage in this sense is a comforting tale in private markets, although some behavioralists doubt its accuracy or at least its breadth. Be that as it may, there is simply no obvious arbitrage mechanism in public finance, where the presence of lawmakers creates a principal-agent problem at the core that we discuss further below. Indeed, public finance can be defined as the study of non-market economic activity. Standard rational choice gives reason to fear the difference: whereas in private markets, arbitrage of the sort described above is a private good, the benefits of which can inure to

---

11 For good general surveys, see Kahneman and Tversky 2000, Baron 2000. Rabin 1998 has been highly influential.
the individual agent, in public finance arbitrage is a public good, the benefits of which inure to the general public. A short seller in a financial market, for example, can capture profits from the irrational exuberance (if such it is) of others; but the political actor who aims to lay bare the illusion of “hidden taxes,” say, can have no assurance that he, personally—or even his political party generally—will benefit from any efficiency gains. Standard rational choice theory predicts that debiasing or arbitrage in the public sector will be undersupplied, increasing the stakes for behavioral public finance.

The parallel also points out something of interest to those who would oppose any “normative” role for behavioral public finance. The questions are obvious: if preferences are inconsistent, how can any lawmaker choose which one is “correct”? What does the consumer sovereignty principle even mean in the face of consumer inconsistency? These are hard questions, to be sure, yet private markets do not ponder them: private markets are relentlessly wealth maximizing. Wealth maximization is a compelling norm in public finance, as well. When two states of affairs can be chosen by subjects depending on their frames or purely formal properties, but one state is more efficient in the sense of producing a higher aggregate value of goods and services, there ought to be a presumption in favor of the more “efficient” frame.

II. Behavioral Public Finance: Three Views of a Possible Cathedral
In this Part, we sketch out three broad clusters of questions raised by behavioral public finance. We mean this discussion to be illustrative, both of the range of issues within each cluster, and of the field of behavioral public finance itself.

A. Form Matters: Framing and Other Optics of Public Finance
A central descriptive component of behavioral economics is that form matters. Contrary to the dictates of ideal rationality, the purely formal aspects of a choice or decision set often affect substantive outcomes. This leads to violations of such basic axioms of rational choice as transitivity and the independence of irrelevant alternatives: preference

---

14 McCaffery and Baron, 2004.
15 Kaplow and Shavell, 2000, Baron and McCaffery 2004.
shifts and reversals can turn on logically irrelevant matters such as how a choice set is described. One cluster of questions for behavioral public finance turns on these violations of the rationality assumption: what implications do citizen heuristics and biases have for central public finance questions?

Consider, for example, the framing effect, under which individuals respond to the purely rhetorical characterization of a constant set of facts, such as preferring a “half full” to a “half empty” glass. An instance of framing relevant to public finance is the metric effect: subjects react differently depending on the unit in which a question is posed, preferring, for example, a tax system featuring higher taxes when asked about taxes in percent rather than in dollar terms. Other examples include penalty aversion, where people prefer policies described as bonuses to their punitive converse (child bonuses versus childless penalties); the Schelling effect, wherein people want progressive bonuses (more for the poor than the rich) and progressive penalties (higher for the rich than the poor), which are inconsistent given a trivial framing manipulation; and tax aversion, wherein people prefer government surcharges described as something other than a “tax,” such as a “payment” or “user fee.”

Real world evidence suggests that successful politicians have at least intuited many of these heuristics and biases.

In the variously described endowment effect, loss aversion, status quo bias, or reference-dependent utilities, subjects react differently depending on their perception of the baseline or status quo: experiencing more disutility from a loss off a high baseline than from a corresponding failure to obtain a gain from a low baseline. The baseline itself can be set arbitrarily—as a matter of framing—and still affect choice. In a classic example from Thaler, individuals will use cash to avoid a penalty for using credit cards at the gas pump, but will foreswear from using cash to obtain a bonus for doing so, on the very

---

16 That form matters in consumer decisions is well understood by marketing directors. Witness the proliferation of cereal boxes that cost $3.99 and gasoline that sells for $1.499 per gallon, and the ubiquity of discounting from “regular” prices. Krishna and Slemrod 2003 address to what extent marketers’ insights can explain income tax design features.
17 McCaffery and Baron, forthcoming; Levin et al., 2002.
18 McCaffery and Baron 2003.
same facts (e.g., $1.90 a gallon for cash, $2.00 for credit cards). In public finance, the endowment effect may lead to a “stickiness” of public goods, such as social security benefits: once in place, citizens will react to their loss more harshly than the failure to obtain an equivalent good. The endowment effect is also one reason why socially set default rules may matter, and why citizens might prefer “fully hidden” taxes to more transparent ones.

Individuals also employ mental accounts; fail to integrate across similar categories, or suffer from an isolation effect. Although money is fungible in the rational model, individuals react as if different sources of wealth map up with different uses: seeing lottery proceeds, for example, as windfalls that need not be devoted to “ordinary” wants and needs. In public finance, this effect can work with the now well-documented flypaper effect as well as the endowment effect to suggest a path dependence to fiscal outcomes. Revenue sources flow to certain particular public uses and “stick” there, making reallocation of funds to higher and better public (or private) uses difficult.

In a related disaggregation effect, individuals have a difficult time integrating parallel but separate systems to form consistent global judgments. Thus, for a major example, it is hard for individuals to take into account the structure of the payroll tax system when making decisions about the appropriate level of progression in the income tax. The same effect suggests that citizens will be hard pressed to understand how changes in one tax, such a negative rate bracket under the income tax, can compensate for the structure of other taxes, such as the absence of a “zero bracket” under the payroll tax system. The disaggregation effect suggests that many smaller taxes can add up to a greater total tax burden, with the same psychic discomfort, as fewer larger taxes. Consider also the possible “privatization” of presently publicly provided goods. The two welfare theorems of neoclassical welfare economics suggest that whether a good or service is provided

---

21 Sunstein and Thaler 2003.
22 McCaffery and Baron 2004.
23 Thaler 1999, McCaffery and Baron 2003.
25 McCaffery and Baron 2003.
26 McCaffery and Baron 2003.
publicly or not, and, if so, how, should be decided by standard welfare maximizing principles. The greater social wealth thereby generated can be redistributed to meet the Paretian constraint. But the disaggregation effect suggests that subjects may not redistribute sufficiently in the tax system, standing alone, to counterbalance the effect of the privatization.\footnote{Baron and McCaffery 2004.} This possibility pits efficiency against equity.

This brief summary overview suggests the stakes for behavioral public finance in terms of these well-noted violations of the rationality assumption. One, as long as the form of public finance matters, politics will turn to at least some extent on formal matters or rhetoric. Politicians will invest time and money in rhetoric, and the better rhetoricians—or salespeople—will have an advantage. Two, and worse, public finance may be unstable, because different frames can elicit preference reversals, the key finding of the Kahneman and Tversky framing literature. As new actors enter the political scene and set new frames, public opinion may shift, supporting a high level of turnover, as it were, in public finance systems, with the attendant transition and transaction costs. Three, and worst of all, real wealth can be left on the table, an homage to cognitive illusions, if the more attractively framed public finance form is not the one that maximizes social welfare. Politicians may choose to please voters with an inefficient tax or spending program; taxes have real effects, independent of their optical or cognitive properties.

It is beyond the scope of the present paper, of course, to sketch out a full political economic model of public finance with behavioral biases in play; such is one of the great potential projects for the field of behavioral public finance. But we can say a few things about what such a project might look like. It is by now well accepted that utility-maximizing applies to politicians as much as to ordinary citizens. Gary Becker (1983) has argued for a model in which politicians reward “pressure groups” with favors (tax breaks, spending programs and so forth), constrained by the opposition of the residual taxpayer class. In the Becker model, equilibrium is reached when efficient interest groups are rewarded by efficient taxes. But what if some taxes produce less psychic pain, strictly on account of their formal properties, than others? Taxes, even “hidden” ones,
have real effects on prices, regardless of their perceived burden. What if politicians are choosing suboptimal taxes, from a welfare maximizing perspective, to make citizens feel better because of the form of the tax?

Much evidence suggests that frequent players in private markets overcome heuristics and biases. Much evidence suggests that frequent players in private markets overcome heuristics and biases. After all, “success” in private markets is measured by the objective benchmark of wealth, and turns on matters of skill. The cognitively disadvantaged is economically disadvantaged, too. In public settings, in contrast, political success turns to a considerable degree on rhetorical skill. It may be that successful policy-makers are indeed subject to the same type of cognitive biases catalogued above. We must consider the subject of behavioral public choice to complement the field of public choice that James Buchanan and others have championed. If not exactly a theory of the blind leading the blind, behavioral public choice must be a theory of the cognitively-biased leading the cognitively-biased, without the same disciplining mechanisms of the private sector to help see us all through the maze.

A related question is, where behavioral heuristics and biases have led to suboptimal public finance structures, whether some form of debiasing mechanism can improve welfare, measured from a more “enlightened” or consistent set of preferences? Suppose, that is, that individuals prefer hidden taxes over direct ones, in part because of loss aversion: they do not notice a loss when they have failed to obtain the gain in the first place. This might for example explain the persistence and popularity of the corporate income tax, which diverts resources from their otherwise ultimate placement in private hands, such that its ultimate incidence is difficult to ascertain. Suppose further that the corporate income tax is more distorting, in the traditional meaning, than alternative sources of revenue—and perhaps suppose, too, that the “hidden” tax does not simply replace other, more transparent taxes, but leads to a larger government. In such a case real wealth is being left on the table, a sacrifice to cognitive illusion and inconsistency. Should public finance help to lay bare the illusions? And, if so, how?

29 McCaffery 1994, McCaffery and Baron 2004.
The aggregate effect of this panoply of as-yet disconnected cognitive biases on the big questions of public finance is not clear. Will government be “too big”? Too slanted towards programs with highly salient, short-term benefits? Too dependent on hidden, excessively distorting taxes? These central questions of public finance should be reexamined in the light of what we are learning about cognitive biases. If the result of the biases is a collection of tax and expenditure program elements that resemble the packaging, price presentation, and product placement one encounters at the local supermarket, then the stakes are not necessarily high. But recall again the disciplining effect of private markets, missing at least in part from government actions. What if, in public finance, Smith’s celebrated invisible hand is replaced with an invisible sleight of hand?

Individuals’ cognitive biases may also affect their behavioral responses to tax and other public finance mechanisms, a central concern of public finance. What behavioral assumptions should lawmakers employ when modeling the effects of tax or other public finance reform? Do for example citizens react differently to something called a “tax” as opposed to something else? Recent experimental work on the “crowding out” hypothesis suggests that they do, reducing their charitable contributions when “taxes” go to the charity, but not when unlabeled exactions do.31 Do the observed labor and capital supply elasticities to tax-law changes reflect behavior biases? And so on. As a matter of applied public finance, we need models that are behaviorally realistic, and the insights of behavioral economics may be indispensable in such settings.

B. Time Matters Too: Time Inconsistency and Problems of Self Control

The first cluster of questions concerned an array of heuristics and biases that call the rationality assumption into question, rewarding rhetoric among politicians, making public finance potentially volatile, and running the risk of leaving real wealth on the table. A second cluster of questions concerns a more specific set of inconsistencies that calls into

the question the very meaning of the consumer sovereignty principle. Many people seem to have inter-temporal preferences that are not only present-oriented but also time-inconsistent. People act as if they do not have the self-control to resist behavior that has short-term benefits but larger long-term costs. There is also considerable evidence that many people misforecast their own future preferences: as two future periods get closer, they give higher relative weight to consumption in the earlier period. One specific form of time inconsistency, known as hyperbolic discounting, has attracted much recent attention.\textsuperscript{32} Under it, the discount factor between consecutive future periods is constant, but is far smaller than the discount factor between the next period and the immediate one, leading to a rolling present-tense bias.

This phenomenon has potentially profound implications for both the positive and normative aspects of public finance. These implications for behavior depend to some extent on whether people are aware that, once the next period arrives, they will become impatient with respect to the new current period and the new next period. If a person is sophisticated enough to realize that she will change her mind in the future, she can make decisions now accordingly. In particular, it may be advantageous to pursue self-commitment devices that limit future choices, like Ulysses tying himself to the mast.

Important applications of related models are to saving and retirement decisions; people seem to save too little on their own, all but certain to later rue the day they failed to save more. This kind of myopia has been used as a justification for a system of compulsory saving like Social Security, and even for specific design features such as the payment of benefits only as annuities.\textsuperscript{33} O’Donoghue and Rabin (1999) explore how time-inconsistent people will tend to procrastinate in preparing for retirement, and suggest that default investment options and imposing deadlines on financial decisions might satisfy a criterion of “cautious paternalism”—valuable for people who are making errors, but with

\textsuperscript{33} Aaron 1999.
relatively small costs for people who are fully rational. Experience has shown that
default rules can have a powerful influence on the saving behavior of employees.34

More recently the model of inter-temporal choice has been applied to addictive and
harmful commodities, defined as those goods for which past consumption increases the
attractiveness of current consumption, and for which future costs are large relative to
present benefits. Imagine that the consumers of such goods are rational and forward-
looking, in the sense of Becker and Murphy (1988), but also have time-inconsistent
inter-temporal preferences and are unable to overcome the self-control problems these
preferences imply. In this setting there is a justification for “sin taxes” that help prevent
present selves from acting for their own future harms. The optimal sin tax can be
calculated using the standard optimal taxation framework, assuming that the policy
objective is to maximize utility based on long-run preferences rather than the preferences
that guide the actual decisions. Gruber and Koszegi (forthcoming) estimate the optimal
tax on cigarettes to be at least $1 per pack, and quite likely much higher.35 Furthermore,
the utility or deadweight cost of a tax on a harmful addictive good is lower than for a
non-addictive good because the consumer places a positive value on the self-control
provided by the higher price. Indeed, Gruber and Koszegi argue that for a wide range of
parameter values, levying a tax on an addictive commodity will on net be beneficial to
the addicted person.

The welfare analysis of optimal cigarette taxes illustrates a very sticky question: where
welfare economics seeks to maximize the satisfaction of individual preferences, as it
typically does, which preferences should it use? The fact that the individual is
“impatient” when faced with a choice between today and tomorrow, but would like to
become patient in the future, creates a conflict between the “current self” and the “future
self.” This type of question is acute where time inconsistency is in play: if policymakers
know that citizens want to spend (smoke) today, but are likely to rue their failure to save

34 Choi, Laibson, Madrian, and Metrick 2004 and Thaler and Bernartzi forthcoming.
35 O’Donoghue and Rabin 2003 carry out a similar exercise.
(quit smoking) tomorrow, which set of preferences—which self—should they choose to please?

Note that the paternalistic argument has moved from arguing that a judicious setting of a default rule can benefit many people while harming none, in the manner of Sunstein and Thaler 2003, to arguing that an inescapable tax on certain behaviors can benefit precisely the people who behave in the targeted way. One may be concerned about high taxes on cigarettes, or on alcohol, unhealthy eating habits, on both vertical and horizontal equity grounds. Take cigarettes. Gruber and Koszegi (2002) present data showing that the share of income spent on cigarettes is eight times as high for the bottom income quartile compared to the top quartile, and is four times as high when the quartiles are defined by consumption, arguably a better indicator of lifetime income. In the context of the self-control problem, the beneficial effects of the tax on the smoker’s own welfare undercuts the apparent regressivity of cigarette taxes.36

Of course, the possibility of widespread time inconsistency looms over one of the most daunting of today’s public finance issues: the implications of future liabilities, as for social security and Medicare, on present government finances.37 A strictly rational choice or rational expectations perspective would suggest that citizens today account for all known future liabilities.38 But do they? And, they do not, are there mechanisms to make future liabilities more salient today? How should one even think about such problems under the consumer sovereign principle, with the problems of this and the preceding cluster of behavioral issues in play? If people want to live for the moment, and let the future care for itself, running up large deficits or whatever in the meanwhile, who are we—who is any “we”—to tell them otherwise?

36 Gruber and Koszegi 2002 conclude that, as long as the poor do not have life values and/or marginal damage from smoking very far below the rich, and as long as their discount rate is not much lower, the regressivity of cigarette taxes is reduced for sophisticated time-inconsistent smokers.
37 Jackson 2003.
38 Barro 1974.
C. Compliance Matters: Towards a Realistic Model of Citizen Duty

The first cluster of questions looked at inconsistencies and confusions in the popular perception of public finance system design; the second raised questions about how to even think about, let alone implement, welfare improving fiscal policies in the light of behavioral inconsistencies. These are questions at least in part of high theory. The final cluster of questions concerns a more practical, applied subject matter for public finance: why do citizens pay taxes, and how can a government keep them doing so/prevent them from not complying? In this area the issue is not only to address challenges to the rationality assumption, but also to understand behavior that appears to violate self-interest, narrowly defined.

Because one’s own benefit from government activity is, with some exceptions, not significantly affected by one’s “contribution,” no purely self-interested person should voluntarily contribute to a public good—that is, pay taxes—unless the threat of punishment makes it sensible. Yet, although “free riding behavior” is indeed widespread, much experimental work (and anecdotal evidence) suggests that free-riding behavior is context-specific. Ostrom (2000, p. 140) remarks that the finding that “the rate of contribution to a public good is affected by various contextual factors” is one of several phenomena that “have been replicated so frequently that these can be considered the core facts that theory needs to explain.” What is going on? There is an active literature that seeks to determine to what extent such behavior is motivated by pure altruism, in the sense that people put positive value on the well-being of others unconditioned by their behavior, or reciprocal altruism, under which preferences over other people’s well-being depend on the behavior, motivation, and intentions of those other people.39 Either one of these behaviors would be “rational,” of course—there is no disputing tastes, after all—although the factual bases for at least reciprocal altruism might not obtain, and any static level of compliance in a model depending on reciprocity is unlikely to be a stable equilibrium. But it is also possible that “excessive” compliance might relate to an irrationality, a behavioral anomaly or inconsistency of some sort. The stakes here can be high: Whatever the motivation, the fact is that the cost of raising taxes, and of running

---

39 The theory and evidence concerning reciprocal altruism is summarized in Fehr and Schmidt 2003.
government, is lower to the extent that taxpayers “volunteer” to comply. This argument applies more broadly than to compliance with the tax system. Christopher Clague (1993, p. 412) argues that “a society with very low levels of rule obedience\(^\text{40}\) cannot…have a net of institutions that is conducive to economic progress.”

The standard public finance model of the demand for tax evasion, due to Allingham and Sandmo (1972), assumes free riding to be ubiquitous. It models the compliance decision as a choice under uncertainty made by amoral individuals, whose decisions depend—strictly—on the chance of being caught and penalized, the penalty imposed, and one’s risk aversion. Looking at real world data, some have argued that this model is flawed: given the probability of audit and the penalties typically assessed, evasion seems to be a winning proposition for many more people than actually do evade. This suggests that something is going on outside of the “standard” rational choice model—either in the utility functions, as in the altruistic and reciprocal altruism explanations, or in some failure of rationality.

It is always important to walk before we run: we ought to get the facts straight. The simple fact is that the over-quick critique of the standard, Allingham-Sandmo style model, is not fully compelling. It is true that the average audit rate for individual tax returns in the United States is less than one percent. With that probability and with the penalty rates in effect, what we know about the degree of risk aversion from other contexts suggests that compliance should be much, much lower than it apparently is, leading to the simple critique. But the flaw in this summary argument is that the one percent probability of detection is a significant understatement of the chance of being caught evading for the bulk of income subject to tax. A wage or salary earner whose employer submits this information electronically to the Internal Revenue Service (as on

\(^{40}\) The use of the term “obedience” will raise a red flag among those familiar with the psychology literature that addresses how people behave toward authority, because what psychologists have learned here is very unsettling. In perhaps the most controversial psychology experiment of all time, Stanley Milgram 1974 demonstrated that, if so instructed by an authority figure, most ordinary citizens would deliver apparently very painful electrical shocks to apparently innocent subjects. That most people are malleable with respect not only to the use of well-intentioned default rules and debiasing, but also to malevolent suggestions, raises concerns about the slippery slope of paternalistic policies. Here we just note this important issue, and leave the debate to the (behavioral) political scientists to pursue.
W-2 forms or 1099’s), but who does not report that income on his own personal return, will be flagged for further scrutiny with a probability much closer to 100 percent than to one percent; much the same can be said for ordinary savers receiving 1099s from their banks and brokers. The same story applies in other countries that have a working system of employer withholding and matching of information returns. Looking exclusively at areas where there is no strong system of third-party reporting—as for the self-employed, or for unregulated asset classes—the rate of compliance seems to be far lower, getting closer to the rational choice model’s predictions.

Still, however, there appears to be some degree of over-compliance, even in these more hidden corners of the economy. Why? Proposed solutions to this puzzle by definition involve pushing beyond the standard economic model, either by enriching it, in the utility functions, or by trying something else altogether. Thus for example, substituting the expected-utility-maximization framework with an alternative framework, in the spirit of Kahneman and Tversky 1979’s “prospect theory,” has immediate implications for the theory of tax evasion. Loss aversion relative to a reference point defined by no evasion will reduce the attractiveness of evasion, because the harm of getting caught will seem worse than the benefit from evading, even if the two are of equal magnitude. Much the same effect can obtain by overweighting the low probabilities of detection and the penalties for evasion.41 Dhami and al-Nowaihi (2004) argue that such a framework (and a stigma cost to for discovered evasion) can more satisfactorily explain the level of observed evasion, the non-ubiquity of evasion, and the fact that tax rates negatively impact evasion.42

41 Bernasconi and Zanardi 2002 explore the implications of reference dependence. Alm, Jackson, and McKee 1992 discuss experimental evidence that is consistent with the latter explanation. See also Yaniv 1999.
42 Dhami and al-Nowaihi’s numerical simulation exercises use parameters based on independent experimental evidence, but strain to explain why observed evasion is so low in light of an assumed probability of detection between one and three percent. As argued above in the text, however, the actual probability of detection for income subject to withholding and verifiable from third-party information returns is much higher than this, so that the expected utility model does not grossly under predict tax evasion of this kind of income, after all.
There are also indications that individuals’ tax compliance behavior depends on variables that lie outside of the free-rider cost-benefit calculus. Some laboratory experiments have found that subjects respond not only to the probabilities and stakes of a tax evasion game, but also to context provided to them, although this finding is not widely documented. Analysis of survey data has yielded mixed results, with an early study concluding that reported compliance is not directly related to dissatisfaction with the tax system, and a more recent study finding that self-reported tax compliance behavior is significantly negatively associated with self-reported trust in government to do what is right and the level of dishonesty in government. Surprisingly and apparently contradictorily, the same study found that political efficacy (whether the respondent has a “say” in what the government does and whether it is run “mainly for the benefit of special interests”) increases noncompliance, perhaps because political efficacy may lead to a perceived ability to manipulate the system without risk.

In their review of tax compliance research, Andreoni, Erard, and Feinstein (1998) identify three classes of explanation for why observed evasion is apparently lower than conventional economic models of tax evasion predict: moral rules or sentiments that determine the psychic costs of evasion, evaluations of the fairness of the tax code and its enforcement, and evaluation of government expenditures and corruption. Frey (1997) links the first two classes of explanation by differentiating between intrinsic and extrinsic motivation. With intrinsic motivation, taxpayers pay because of “civic virtue;” with extrinsic motivation, they do so because of threat of punishment. Frey argues that increasing extrinsic motivation—say with more punitive enforcement policies—“crowds out” intrinsic motivation by making people feel that they pay taxes because they have to, rather than because they want to. Similarly, in Cullis and Lewis (1997), individuals care not only about their own consumption, but also value their own compliance with the

---

43 Spicer and Becker 1980 and Alm, Jackson, and McKee 1992 find support for this proposition. Cowell (1990, p. 219) reports on experiments that fail to find links between perceived inequities in the tax system and non-compliance.

44 Mason and Calvin 1984.

45 Scholz and Lubell 1998.

46 Scholz and Lubell 2001, in an experimental setting, find that the level of cooperation in certain settings declines significantly when penalties are introduced, suggesting that the increased deterrence motivation did not compensate for the change in decision frame brought about by the penalties.
social convention of tax compliance and separately the extent of others’ compliance with the norm, either directly or indirectly via pecuniary consequences. If tax equity strengthens the social norm against evasion, then evasion becomes more costly in terms of bad conscience (if not caught) or bad reputation (if caught) in a society with a more equitable system. An individual can also find unfairness in goods provision due to the provision of the wrong goods—someone such as Thoreau may avoid taxes because he thinks government policy wrong. But this is not a simple matter. Expenditures on warfare might be tolerated in a patriotic period but rejected during another period characterized by anti-militarism. Expenditure on welfare might at times be seen as a socially desirable pooling of risk, and be seen at other times as a source of national decay. And so on.

All this behavior suggest that reciprocal altruism may be at work, but where the taxpayer’s behavior depends on the behavior, motivations, and intentions not of any subset of particular individuals, but of the government itself. Some taxpayers may be willing to give up some of their own expected utility by failing to optimize their evasion, in order to effect a fairer distribution of outcomes, but only if they perceive the tax system and tax enforcement process to be fair. This characterization is very similar to the spirit of Levi (1988, p. 91), who argues that citizens are likely to trust government only to the extent that they believe that it will act in their interests, that its procedures are fair, and that their trust of the state and others is reciprocated. Moreover, government trustworthiness, plus the perception that others are doing their share, can induce people to become “contingent consenters” who cooperate even when their short-term material self interest would make free riding the individual’s best option. In Levi’s words, “the willingness to pay taxes quasi-voluntarily or to give one’s contingent consent to conscription often rests on the existence of the state’s capacity and demonstrated readiness to secure the compliance of the otherwise noncompliant.” Once again, this

47 Falkinger 1995.
48 Daunton 1998 makes this argument.
49 Levi 1997 reminds us that military service is another important way that democratic governments are able to elicit both money and men from their populations in the face of tax evasion, draft evasion, and other forms of disobedience.
can be a strictly “rational” matter, a result of preferences, or the government’s activities can make the penalties more salient, exacerbating a prospect-theory like effect.

What does this imply about optimal tax enforcement policy? Feld and Frey (2002) argue that to sustain citizens’ commitment to the contract and therefore their morale, the tax authority must act respectfully toward citizens while at the same time protecting the honest from the free rider. It does this by giving taxpayers the benefit of the doubt when it finds a mistake, by sanctioning small violations more mildly, and by sanctioning large and basic violations (e.g., the failure to file a return) more heavily.

There is little evidence outside of the laboratory that such policies affect noncompliance conditional on their deterrence effects. The finding of Feld and Frey (2002) that blatant violations of the tax code are punished more heavily in the more direct democracies among Swiss cantons provides some evidence that citizens in more direct democracies are not simply pushing for leniency but rather evaluating behavior as upholding or violating a basic contract. Empirical documentation of the determinants of contingent compliance by taxpayers is, however, sparse. Just sending a letter reminding taxpayers of their civic duty will not for example affect compliance, as demonstrated by Blumenthal, Christian, and Slemrod (2001).

It is notable that all of the literature about whether attitudes affect compliance applies to individual taxpayers, although in most countries the bulk of taxes are remitted (as opposed to borne, in the sense of ultimate incidence) by businesses, either because the taxes are levied on business entities or because labor income taxes are withheld by the employer. Whether a company’s policy would react as an individual is a fascinating and completely open question, one that is related to the motivations behind corporate charitable contributions. This query applies more broadly—under what circumstances do organizations mitigate or perhaps exacerbate cognitive biases?

50
III. Objections and Conclusions

We cannot conclude without at least noting that there remain skeptics to the whole enterprise of behavioral economics. Some still deny that the various heuristics and biases exist, although this is getting increasingly hard to do in the face of abundant experimental and real-world evidence. The more enduring critiques fall into two major (and somewhat related) camps. One holds that the biases might exist, but they are artifacts of the experimental design or other institutional setting in which they are found.\textsuperscript{51} Better design, education, incentives and so on can mitigate or altogether eliminate the biases.\textsuperscript{52}

A second camp picks up the theme, of Barberis and Thaler and others, of arbitrage mechanisms. Here, the reasoning goes, conceding that heuristics and biases exist and are even rampant—that individuals cannot and do not overcome them (individual-level debiasing, as through education, has its limits)—the biases do not much matter, because they do not materially affect how efficiently markets work. Systems solve individual errors. Thus for example behavioral economics may not pose a challenge to perhaps the most celebrated finding of the standard rational choice view, the efficient market hypothesis: even if almost all agents are irrational, markets can still work, in a fashion to do Adam Smith proud.

We have no pony in this now familiar race. What we find interesting is that the by-now standard responses to the perceived challenge of behavioral economics in private market domains depend on factors—learning, learning by doing, incentives, arbitrage mechanisms such as the market or the impersonal forces of competition—that may be altogether lacking and in any event are very different in the public setting. Inefficient structures ought not to persist long in private markets, as long as they are at least some one-eyed persons in the land of the blind. But unless the one-eyed is also a saint, public finance has a long way to go.

\textsuperscript{50} Arlen, Spitzer and Talley 2002 find some support for the idea that the endowment effect—the tendency to value goods more highly when one perceives an “ownership” of them—is lessened in the agency context familiar to the corporate world.

\textsuperscript{51} Plott and Zeiler 2002.

\textsuperscript{52} List 2004.
This leads us to continue the quest. But we caution once again against running before walking. Behavioral public finance is too new, and too little time has spent studying behavioral biases in the public context, to recommend any radical changes to the traditional practice of public finance research. Instead, we more humbly suggest that the research agenda we hope to have helped begin, here and in two related conferences, continues. More research is needed on the heuristics and biases in the public sector. Models of behavioral public choice need development. Searches for arbitrage mechanisms in the public sphere should begin. Behavioral insights should be employed to help model and solve some longstanding puzzles of public finance, such as what appears to be over-compliance.

In the end, we may only be left with better questions, and few answers. But that’s a start.
References


