

Introduction: Symposium on *Escaping Paternalism: Rationality, Behavioral Economics and Public Policy* by Mario J. Rizzo and Glen Whitman

Nick Cowen¹ and Malte Dold²

¹*School of Social and Political Sciences, University of Lincoln, Lincoln, UK;*
ncowen@lincoln.ac.uk

²*Economics Department, Pomona College, Claremont, CA, USA;*
Malte.Dold@pomona.edu

ABSTRACT

Rizzo and Whitman’s *Escaping Paternalism* (2019) is, at once, a scholarly treatise on the nature of rationality and a powerful critique of the use of behavioral insights to support a new paternalism in public policy. Since its recent publication, it has informed research, among other things, into the decision processes of paternalist policymakers (Ambuehl *et al.*, 2021), the implications of dynamic preferences for tax policy (Delmotte and Dold, 2021), and alternative theoretical grounds for behavioral policymaking (Sunstein, 2021; Sugden, 2021; Rizzo and Whitman, 2021). Its theoretical depth has far-reaching implications for methodological discussions within behavioral economics and the scope of government action beyond contemporary policy debates. We are very grateful to the editor of the *Review of Behavioral Economics* for hosting this critical interdisciplinary discussion of the book. In this introduction, we briefly review the key arguments from *Escaping Paternalism* and then summarize the contributions to the symposium. The diverse views expressed therein show that Rizzo and Whitman’s critique of the methodological robustness of behavioral biases and their proposal of a “paternalism-resisting framework” is stimulating, but not uncontroversial. This special issue is an invitation to further engage with Rizzo and Whitman’s arguments and, in doing so, advance both the methodological debate about key concepts in behavioral economics (such as rationality, biases,

internalities, or welfare) and the normative debate about the implications of behavioral insights for policymaking.

The Book

Escaping Paternalism begins with a discussion of the fundamental principles of economic rationality, transitions into a meticulous analysis of the evidence of bias in individual decision-making, and ends by turning the critical gaze on regulators, pointing out where their behavioral biases are likely to emerge. Chapter one establishes the key target of this book, the new paternalism that, unlike the old paternalism, aims to improve people's conduct according to their own lights rather than an imputed objective notion of welfare or moral standards. It characterizes the book as presenting a gauntlet of challenges that proponents of the new paternalism must pass. Failure at any step means that the overall case for these interventions, which include sin taxes, default-choice design and using graphic images and other prompts to influence purchasing decisions, stands on shaky grounds.

Chapter two defends the notion that humans are rational agents by introducing the concept of *inclusive rationality*, a rationality that involves purposeful action plausibly aimed at achieving goals but is comfortable with realistic limitations on knowledge and cognition. Rizzo and Whitman explain that when facing complex decisions, the widespread use of heuristics may turn out to be cognitively efficient. In chapter three, this position is contrasted with the neoclassical "rationality for puppets" that insists that, to be rational, agents must display consistent, that is well-defined and transitive, preferences. Rizzo and Whitman explain that this conception of rationality was not initially intended to act as a normative baseline but rather as a descriptive model of how economizing agents behave in certain competitive conditions. There is no reason to believe that holding consistent preferences improves welfare, only that occasionally (as in the case of a money pump scenario) a specific inconsistency might need to be resolved.

In chapter four, Rizzo and Whitman critically assess the evidence for some key examples of these preference inconsistencies. They find that supposed time-discounting inconsistencies can be explained by humans generally experiencing time in a non-linear fashion; that the experimental evidence for endowment effects are surprisingly weak and that its occasional appearance can be explained by a rational aversion to loss; and that supposed overestimation biases in our capacity to predict our future emotional states are often helpful for agent learning. In all cases, human behavior is too complex to characterize simply as irrationally "biased" and it is unclear what the normative basis is for correcting

these “biases.” Chapter five turns the critique from preference biases towards biases in beliefs. Economists generally assume that useful beliefs are meant to be truth-tracking, and therefore should be logically coherent and responsive to new evidence. Rizzo and Whitman challenge this normative assumption. Biased beliefs might be important sources of personal motivation or function as a source of pleasure or comfort in themselves. Holding onto such beliefs can be perfectly rational (in an inclusive sense). Even when it might be useful, holding a perfectly logical set of beliefs might be too cognitively demanding. For instance, when updating the probability of events or outcomes in light of new evidence, agents may rationally weight their own local experience more strongly than formal base rates. Or, when receiving factual information, agents may rationally include tacit knowledge from the context and implication from the speaker that would not be captured in a formal account of evidence.

In the second half of the book, Rizzo and Whitman examine what this critical stance on the heuristics and biases literature implies for behavioral public policies. In chapter six, they argue that there are reasons to doubt the results of experimental research will translate into real-world settings. Experimental findings can be highly influenced by the artificial choice context in the computer laboratory, which means that estimating the relevant impact of any cognitive bias would similarly have to be based on the real-world context in which a proposed policy is to be implemented. Many behavioral economic experiments involve hypothetical choices, and few of the results drawn from such experiments are reproducible in real-world scenarios. Real-world experience presents agents with many opportunities to learn. For example, incentives that reward accuracy can reduce or eliminate endowment effects over time. Unlike in laboratory experiments, many important decisions in the real-world are taken in small groups, that are known to perform better than individuals.

Chapter seven radicalizes this point with an argument that policymakers face an intractable knowledge problem when attempting to fine-tune their interventions to improve individual decision-making, a problem that parallels historical attempts by socialist governments to plan economies. Policymakers cannot easily know the “true” preferences of specific individuals impacted by their proposed interventions nor the extent of their biases. They typically do not know what steps individuals have taken spontaneously to manage their biases, such as reference to group decision-making, and self-regulatory behavior that might be crowded out by greater external control. In addition, while in experimental settings biases are measured in isolation, in a more complex real-world scenario multiple biases can operate in conjunction, potentially offsetting or correcting for each other. Interventions generally aim to improve an average outcome without accounting for the preference and bias heterogeneity within the population.

Chapter eight turns to the constraints placed on behavioral public policies from a public choice perspective. Rizzo and Whitman warn against the

“nirvana fallacy,” the idea that observing apparently sub-optimal decisions in civil society is sufficient to justify policy intervention. The policy interventions might themselves be sub-optimal and potentially worse than what individuals engaged in learning and adaptation may come up with spontaneously. Rizzo and Whitman argue that the case for intervention should proceed from an assumption of behavioral symmetry. If people generally have cognitive biases, then policymakers (also being people!) will plausibly act on a similar set of cognitive biases. Indeed, one sees examples of overconfidence, confirmation, and salience biases all the time in justifications for real-world policies. Even if policymakers engage in deliberative and reflective decision processes, in a democratic society they will be under pressure to respect the cognitive biases of their constituents. Alternatively, if they are insulated from voters, policymakers may be incentivized to enact rules that align with special interests, such as actors in business or civil society who can profit from new regulation. Seldom will they be motivated to enact rules that harm their own interests, such as those that reduce their influence even if they are the best option for consumers or citizens.

Chapter nine addresses a defense of the new paternalism, that it is essentially moderate in character, with an argument that even light-touch paternalist policies have their own momentum through altering the choices and opportunities of policymakers. Paternalist policies are subject to slippery slopes where the vague aims and methods for justifying them mean that the implementation of one policy intervention can make subsequent extensions and enhancements attractive even if they were considered beyond the pale before. The original interventionist policies, and their often-reconstructed rationales, provide a justification or precedent for further, more interventionist, policies. Enacting interventions can create specific constituencies of experts and lobbyists who will press for expansion and resist any reduction in the scope or strength of a policy.

The final chapter sums up the critique of the new paternalism, reminding readers that there are high bars – conceptual, empirical and practical – to successfully improving individual welfare through behavioral interventions, and that a more realistic conception of *inclusive rationality* should replace the “puppet rationality” of neoclassical economics. Rizzo and Whitman note that behavioral economists (of a paternalist predilection) are often attracted to making arguments in favor of enacting specific policies in a way that is dissonant with the complex subject matter of the research that normally only offers tentative results. Instead, Rizzo and Whitman believe that behavioral economists are better situated to advise the public on better decision-making rather than addressing policymakers. They consider some possible responses from behavioral paternalists, including the abandonment of subjective preferences and a greater insistence on non-coercive interventions, but insist that each of these approaches still end up projecting the economists preferences onto others.

The Symposium

Beginning the symposium, Peart (2021) discusses the historical links between the old and new paternalisms, as well as the Millian skepticism that inspires resistance. Rizzo and Whitman align their argument normatively with the Mill of *On Liberty*. Peart shows that Mill's anti-paternalism is grounded in his conception of empirical social science as dealing constantly with "a multiplicity of causes" when attempting to explain any phenomena. On this account, it is difficult to derive anything but general principles of economic behavior and that, therefore, improvement in decision-making cannot be implemented at the top but rather happens through learning, including the experience and discussion of particular cases. She contrasts this with Jevons, a confident proponent of interventionist social reform, who believed that many of the causes of poverty were due to the bad saving and consumption habits of the poor that could be systematically corrected. Jevons used an early neoclassical model of society as a unitary consumer to generate a baseline of consumption that optimized utility across time. Naturally, real consumers fail to follow this ideal. Jevons is confident that he can improve social welfare with education and instruction that brings the decisions of the poverty stricken in line with – with what he sees – their best considered interests. Peart notes that the new paternalists show much more humility than the old paternalism about understanding individual interests. Nevertheless, there is still resemblance in some of their methods.

Matson and Dold (2021) also take a historical perspective and connect Rizzo and Whitman's arguments with ideas from David Hume's essays on happiness. In those essays, Hume illustrates by means of a literary dialogue that there is no one path to happiness, but people have different views of what constitutes the good life. Crucially, these diverging and sometimes conflictual views should not be overwritten by the philosopher-economist. Matson and Dold argue that Hume's open-ended view of welfare are congenial to Rizzo and Whitman's notion of inclusive rationality. They discuss the ways in which Hume's essays lend support to Rizzo and Whitman's plea that contemporary behavioral economists ought to view themselves as friendly advisors in a public, open-ended conversation with fellow citizens, but not as expert advisors to policymakers.

Rajagopalan (2021) continues the discussion of Rizzo and Whitman's concept of inclusive rationality, which she agrees to be a promising framework for modelling human behavior but one with extant gaps in explanation, especially where the limits of rationality are on this account. She contributes to the development of the framework with an interdisciplinary account of internal struggle and aspiration, two connected and familiar innately human sentiments, that are typically treated as anomalies (or assumed away) within the neoclassical framework. People's attempts to improve themselves rarely

take place in a linear fashion, especially in the case of overcoming addictions. Rather improvement in lifestyles happen through repeated attempts at behavioral change and frequent lapses. Yet, the struggle is often worthwhile. As a result, acting on inconsistent preferences may not simply be more efficient than seeking logical coherence in some cases, but be part of a dynamic process of personal re-constitution.

In his contribution, Grüne-Yanoff (2021) takes a more critical stance on Rizzo and Whitman's notion of inclusive rationality. He argues that the notion leads to the *Panglossian conclusion* that whatever individuals do must be best for them. In other words, the notion of inclusive rationality does not provide the conceptual basis for identifying welfare-relevant decision mistakes. In the first part of the article, Grüne-Yanoff challenges Rizzo and Whitman's notion of inclusive rationality and illustrates that many intuitive notions of decision errors require consistency-based principles. In the second part, Grüne-Yanoff questions Rizzo and Whitman's *fatalistic conclusion* that there are no justified paternalistic interventions into individual decision-making. He argues in favor of so-called *boosts* that are epistemically less demanding than the behavioral paternalistic interventions Rizzo and Whitman discuss in their book.

Hands (2021) picks up and expands on Rizzo and Whitman's discussion of knowledge problems in the context of libertarian paternalism. The standard justification for nudge interventions are individuals' decision errors that cause internalities, that is, unaccounted for costs individuals impose on their future selves. Hands largely agrees with Rizzo and Whitman about the knowledge problems that behavioral paternalists face to identify and remedy those internalities. But he also makes clear that a critique of behavioral paternalism does not necessarily lead into a general critique of microeconomic-based governmental policy. Hands defends governmental attention in the area of social policies that economists have traditionally considered to be the responsibility of democratic governments. Hands's contribution is a warning for economists against spending too much intellectual effort on the level of individual preference satisfaction, while losing sight of more pressing social issues, such as externality problems.

Hargreaves Heap (2021) agrees with Rizzo and Whitman's criticisms of the neoclassical approach to modelling economic behavior but believes their approach is orthogonal to the real issues facing policymakers. Politicians are not interested in whether policies have "paternalist" intentions or not, rather whether a policy can help solve whatever problems the public agenda throws up, whether these problems result from internalities or externalities. The nudges that behavioral economists propose can also be used to discourage behavior that impacts on third parties. Yet, when addressing such proposals, Rizzo and Whitman are still highly skeptical of both their efficacy and wisdom because of concerns with efficacy and slippery slopes. This suggests it is not paternalism as such but some other feature that Rizzo and Whitman ultimately take to be

problematic. Hargreaves Heap believes this feature to be a departure from the setting of general rules designed to shape the available choice set for all people and instead an attempt to manipulate what specific individuals choose within a given set of opportunities.

Cowen and Trantidis (2021) also take on Rizzo and Whitman's slippery slope argument and offer a classical liberal defense of what they term "soft interventionism." They agree that neoclassical assumptions of formal rationality offer a poor practical and normative baseline from which to judge individual decisions or justify policy interventions. However, they argue that, from a Hayekian perspective, people rely on an enormous range of cues and default options present in the social environment, some of which can mislead people as to their long-term interests. While individuals and communities are generally successful at noticing and correcting these errors, the state may have a role too so long as these interventions are made in a public and transparent fashion. Cowen and Trantidis point to the example of medicine package design as a way of reducing suicide and accidental overdose. Moreover, the strict division between internalities and externalities is sometimes blurred in practice. In these contexts, interventions short of coercion and restriction can help people better achieve what are jointly public and private goals.

Koppl (2021) characterizes Rizzo and Whitman's examination of biases in behavioral public policy as "foundational" and shows how it forms the basis of a more general analysis of "expertism." Drawing on the evolutionary origins of human conduct, Koppl observes that individuals are often granted status based on their skills that are perceived to be useful to group success. Therefore, individuals are incentivized to publicly display their knowledge and skills regardless of their underlying practicality to exert power over the group. This explains the enduring attraction of paternalism in its various guises. He argues that the professional and political institutions associated with science often seek status and the augmentation of expert authority rather than the protection of scientific inquiry itself. Koppl thus offers a novel, plausibly more realistic, standpoint from which to assess the intent of choice architects.

References

- Ambuehl, S., B. D. Bernheim, and A. Ockenfels. 2021. "What motivates paternalism? An experimental study". *American Economic Review*. 111(3): 787–830. DOI: 10.1257/aer.20191039.
- Cowen, N. and A. Trantidis. 2021. "Soft interventionism: A Hayekian alternative to libertarian paternalism". *Review of Behavioral Economics*. 8(3–4): 341–360.

- Delmotte, C. and M. Dold. 2021. "Dynamic preferences and the behavioral case against sin taxes". *Constitutional Political Economy*. March. DOI: 10.1007/s10602-021-09328-8.
- Grüne-Yanoff, T. 2021. "Boosts: A remedy for Rizzo and Whitman's Panglossian fatalism". *Review of Behavioral Economics*. 8(3-4): 285-303.
- Hands, D. W. 2021. "Libertarian paternalism: Making rational fools". *Review of Behavioral Economics*. 8(3-4): 305-326.
- Hargreaves Heap, S. P. 2021. "The 'problem' is different and so is the 'solution'". *Review of Behavioral Economics*. 8(3-4): 327-340.
- Koppl, R. 2021. "Against expertism". *Review of Behavioral Economics*. 8(3-4): 361-377.
- Matson, E. W. and M. Dold. 2021. "The behavioral welfare economist in society: Considerations from David Hume". *Review of Behavioral Economics*. 8(3-4): 239-258.
- Peart, S. J. 2021. "On making and remaking ourselves and others: Mill to Jevons and beyond on rationality, learning and paternalism". *Review of Behavioral Economics*. 8(3-4): 221-237.
- Rajagopalan, S. 2021. "Inclusive rationality: Struggle and aspiration". *Review of Behavioral Economics*. 8(3-4): 259-283.
- Rizzo, M. J. and G. Whitman. 2019. *Escaping Paternalism: Rationality, Behavioral Economics, and Public Policy*. Cambridge Studies in Economics, Choice, and Society. Cambridge, United Kingdom; New York, NY: Cambridge University Press.
- Rizzo, M. J. and G. Whitman. 2021. "The unsolved Hayekian knowledge problem in behavioral economics". *Behavioural Public Policy*. May: 1-13. DOI: 10.1017/bpp.2021.18.
- Sugden, R. 2021. "How Hayekian is Sunstein's behavioral economics?" *Behavioural Public Policy*. May: 1-10. DOI: 10.1017/bpp.2021.11.
- Sunstein, C. R. 2021. "Hayekian behavioral economics". *Behavioural Public Policy*. March: 1-19. DOI: 10.1017/bpp.2021.3.