Brazilian Journal of Political Economy, vol. 42, nº 2, pp. 532-549, April-June/2022

Does Behavioral Economics substitute or complement Neoclassical Economics? Rethinking the behavioral revolution from a contextualist approach

A Economia Comportamental substitui ou complementa a Economia Neoclássica? Repensando a revolução comportamental por uma abordagem contextualista

JOSAFAT IVAN HERNANDEZ-CERVANTES*

RESUMO: A ascensão da Economia Comportamental levanta uma questão importante: ela substitui ou complementa a Economia Neoclássica? Este artigo faz uma análise para identificar como a questão da substituição-complementaridade é altamente sensível ao contexto epistêmico. Em particular, o artigo distingue quatro contextos epistemológicos: descritivo, explicativo, preditivo e prescritivo. Defende a ideia de que a abordagem comportamental substitua a neoclássica em contextos descritivos e explicativos; no entanto, há uma complementaridade nos contextos preditivo e prescritivo, onde existem alguns domínios em que a abordagem neoclássica ainda pode funcionar bem.
PALAVRAS-CHAVE: Economia Comportamental; Economia Neoclássica; revolução comportamental; contexto epistêmico; racionalidade limitada.

ABSTRACT: The rise of Behavioral Economics poses an important question: does it substitute or complement Neoclassical Economics? This article makes an analysis to identify how the substitution-complementarity issue is highly sensitive to the epistemic context. In particular, the paper distinguishes four epistemological contexts: descriptive, explicative, predictive, and prescriptive. The idea that the behavioral substitute the neoclassical approach in descriptive and explicative contexts is defended; however, there is a complementarity in the predictive and prescriptive contexts, where there are some domains in which the neoclassical approach might still work well.
KEYWORDS: Behavioral Economics; Neoclassical Economics; behavioral revolution; epistemic context; bounded rationality.
JEL Classification: A12; B40; B41; D01; D91.

* Postdoc at the Center for Complexity Sciences – UNAM – Universidad Nacional Autónoma de México, México. E-mail: josafativan.hernandez@gmail.com. Orcid: https://orcid.org/0000-0002-8077-5900.
Submitted: 16/November/2020; Approved: 14/April/2021.

 http://dx.doi.org/10.1590/0101-31572022-3248

INTRODUCTION

Behavioral Economics (BE) has emerged as a critique of Neoclassical Economics (NE). Since the works of Herbert Simon (1955, 1956), increasing literature with empirical evidence has shown that agents are not very rational. They misbehave (Thaler, 2015) because of their bounded rationality, heuristic reasoning, cognitive biases, and emotions (Tversky & Kahnemnan, 1974; Kahneman, 2003, 2011). These cognitive elements do not allow agents to make optimal decisions, which contradicts the NEs assumption of full rationality. This situation is forcing a change in economic thinking where a new notion of agency might emerge. As some authors have suggested, a possible synthesis is underway (Aumann, 2019). But the discussion is whether this theoretical shift will be more neoclassical (Aumann, 2019; Chetty, 2015) or more behavioral (Angner, 2019) in nature.

In this article, I will present arguments in favor of a more behavioral synthesis between BE and NE. But the synthesis is not easy, because the relation between BE and NE is complex. Given the plurality of approaches and subdisciplines in economics, we might focus the analysis on the different epistemic contexts that economists face in their scientific practices. The article, hence, will address the following question: In which specific contexts might BE substitute NE and in which contexts might BE and NE be complementary? There are different ways of understanding what is an epistemic context. In this paper, I will assume a position very similar to that posed by Sandra Mitchell (2009) where we should realize that a model is not always right, but neither is it always wrong. It is context-dependent (Mäki, 2018; Mitchell, 2009). "Nothing is a model in itself. Modelhood requires a larger structure with which an object becomes a model" (Mäki, 2018, p. 4), where some components of epistemic contexts are the following: a) agents who make and use models, b) different pragmatic objectives, where the theory or a model might seek the realization of one in particular, c) different levels of analysis or abstraction, d) an audience with certain characteristics, and e) available research methods and techniques.

The substitution-complementarity issue between BE and NE, hence, might be addressed from a contextual approach, where the main thesis of the article is the following: The neoclassical notion of agent could be substituted or complemented by BE just in relation to the specific epistemic context that economists face in their respective scientific practices. These contexts might be distinguished mainly by the epistemic goals economists try to achieve and, secondly, by the level of abstraction of their analysis[1]. As we will see in the second section, the epistemic goals might be the following: describing, explaining, forecasting, and prescribing. In this article, I will present some reasons to say that BE substitutes NE in describing and explaining human behavior (third section), but not necessarily in forecasting and prescribing it (fourth section). In these two contexts, a complementarity between BE and NE might

[1] The other elements of a epistemic context are relevant, but for the purpose of this article I will focus only on these two elements to make de argument clearer.

be possible because both work at different levels of abstraction. This distinction makes possible a complementary use between BE and NE to improve the evaluation of different scenarios and expand our tools for the design of public policies to address complex social phenomena. I end the paper with some final remarks.

THE MULTIPLE EPISTEMIC GOALS OF ECONOMIC MODELS

Julian Reiss, following the Austrian economist Carl Menger, mentioned that some epistemic goals of economists are explanation, prediction, and control over the economy for the design of public policies (Reiss, 2013). However, as Reiss suggested, these aims are not the only ones. There are others: accurate descriptions of facts and the development of normative stances (Reiss, 2013). In this paper, I will distinguish four epistemic goals that are presented in economists' scientific practices: description, explanation, forecasting, and prescription. This distinction is not exhaustive, but they allow us to develop our argument.

In economics, description focuses on characterizing facts. There are discussions about measures, indicators, and statistics that better describe phenomena (Reiss, 2013). A good description tells us what is happening, what are the main features of phenomena, which agents are engaged in a situation, how agents are related among them and, which specific decision-making problems they face. At a behavioral level, descriptions focus on addressing the following questions: How do agents make decisions, what behavioral patterns do they follow, which specific problems do they have, how do they solve them and, what resources do they have at their disposition to act.

Once the description has been made, the next step is the elaboration of scientific explanations that address why questions, which are answered through hypothesis, models and narratives (Reiss, 2013). A good explanation in economics:

> "[...] will, therefore, ask for the detailed causal process or mechanism that is responsible for the phenomenon of interest [...] a mere regularity that connects an input and an output – no matter how stable and lawlike that regularity is – does not explain why there is a connection between input and output. To investigate a mechanism means to open the black box between input and output and to illuminate why regularities hold and outcomes happen (Reiss, 2013, p. 35).

As we can see, explanation involves the exposure of causal relations, processes or mechanisms that structure phenomena. Scientific explanations in social sciences show which causes are the most relevant, how they are interrelated, and why a phenomenon happens in one specific way and not in another. Modeling here plays a very important role because models help us study a mechanism or causal relations that structure phenomena in isolation (Mäki, 1992, 2018). Idealizations

here are isolating tools that can help economists simplify complexity, leaving aside irrelevant factors that distort the causal relations to be analyzed (Mäki, 1992, 2018).

Even so, we should realize that not all economists explain by causal mechanisms. Some of them might consider valuable another manner of explaining such as unification or hermeneutical interpretations (Reiss, 2013). Moreover, some economists might prefer explaining social phenomena only attending a macro level, a micro level, or a combination of both (Marchionni, 2008). Such explanations might be breadth (unifying a wide range of behaviors) or deep (showing causal mechanisms of singular phenomenon) where different kinds of explanations might be weak or strong complementarity to explain complex phenomena (Marchionni, 2008).

The main task of social sciences is to explain social phenomena (Hausman, 1992; Lawson, 1997). The other epistemic goals, such as forecasting and the design of public policies are dependent on or subordinate to explanation (Lawson, 1997). But explaining is not the only epistemic goal, as we have already said.

Prediction is very important in economics. However, it is problematic because, though historically NE has aspired to have a similar level of accurateness as Newtonian physics, different philosophers of economics (Hausman, 1992; Lawson, 1997) have shown good reasons to say that economics could not have accurate predictions like astronomy. In economics is not possible to predict when and how an economic crisis will happen, the exact rates of economic growth that a country will have in a decade, the accurate exchange rate of a currency over the next year and, when a pandemic such as Covid-19 will emerge and generate a tremendous impact in the global economy.

However, economics could make forecasts. Forecasts answer the question of what would happen if certain conditions remain the same and what would happen if those conditions are changed (Morgan, 2014). Here, once again, modeling is relevant because it allows us to foresee possible scenarios and evaluate in a counterfactual manner the consequences of different actions.

Finally, the prescriptive context tells us how things or a situation ought to be, according to some normative values. Prescriptions tell us what we should do to improve a situation and what we should not do to avoid making a situation worse. According to Reiss, economists discuss how we should characterize normative concepts such as welfare, rationality, and distributive justice (Reiss, 2013). At this level, we could hold discussions about what is efficiency, how we should understand optimization, definitions of economic development, and the role that the state should play in the economy (Reiss, 2013). Normative aspects are present when evaluations of public policies are done, where it is often discussed whether the aims sought after are reasonable, adequate, or viable.

As we can see, economics is a science with multiple epistemic goals (such as the one mentioned above), where the adequateness of methods, techniques, and theories depends on the goals that are pursued (Reiss, 2013). But also, the levels of abstraction, understood as different levels of specificity, are important to make "pragmatic choices about the representations we fashion to deal with that reality.

Different representations and different levels of specificity work for different purposes" (Mitchell, 2009, p. 115).

The next section will show how the substitution or complementarity of BE and NE depends on the epistemic context, where economists seek different epistemic goals.

EPISTEMOLOGICAL CONTEXTS WHERE BE SUBSTITUTES NE

Describing human behavior

Description, as we saw in the previous section, is an epistemic goal of the scientific practices of economists. One of the main criticisms that have legitimated the rise of BE is that the NE model of expected utility does not describe actual human behavior well. Simon, Kahneman, and Tversky have emphasized that agents systematically deviate from the predictions of NE when making decisions. In the case of Simon, he used to say that neoclassical models, particularly the notion of optimization, addressed the question of what agents ought to do, but did not describe how agents make decisions, i.e., NE is normative, not descriptive (Simon, 1955).

Agents do not process the full information they have at their disposition. The time to choose is limited, and in practical cases, agents usually are not interested in making the best decision (Simon, 1955). Especially in changing scenarios where they are pressed by the environment to make a choice. In such contexts of uncertainty, agents only care about making viable decisions, good enough decisions, not optimal decisions.

According to Simon:

> [...] I shall assume that the concept of "economic man" (and, I might add, of his brother "administrative man") is in need of *fairly drastic revision* [...] the task is *to replace* the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms, including man (Simon, 1955, p. 99; the emphasis is mine).

From the above quote we can say that, for Simon, the task is to substitute NE for other models of rationality with more realistic or sound psychological assumptions, i.e., assumptions that are validated by empirical evidence (Simon, 1955). With this idea in mind, Simon developed the notion of bounded rationality, where it is assumed that agents are not fully rational. Agents, on the contrary, have limited cognitive capacities: limited skills for calculation, for information processing, and memory (Simon, 1955).

Simon built an alternative model known as satisficing (Simon, 1955). This model assumes that agents have different aspiration levels and partially ordered preferences where an option has only two values: either it is acceptable, or it is not. Agents also have the cognitive skill of identifying cues in the environment which lead them to satisfy their aspiration levels (Simon, 1955). In this context, agents are

looking only for good enough decisions, i.e. decisions that are above the agents' aspiration level (Simon, 1955). The higher the aspiration level is, the harder a decision will be. And the opposite is also the case: the lower the aspiration level is, the easier a decision will be. Any option that is above the aspiration level is a good enough option. Any option that is below the aspiration level is not a good enough option (Simon, 1955). Mathematical optimization here is not required because agents are not concerned with choosing that (usually unique) option that maximizes their utility function.

Satisficing is a model that describes the actual decision making of real agents situated in real contexts, better than NE does. An example is the following. Let us consider the case of a person who goes to the market, a street market that in Mexico we call "Tianguis". This kind of market is huge. Thousands of people offer and demand several *commodities* at different prices and quantities, one or two days per week in very large streets that they occupy. Let us consider the case of only one person who goes to the tianguis to buy ingredients to prepare a traditional dish. This person faces a variety of *commodities* that have different price-quality relations. How does this person choose from amongst all the options he has at his disposal, considering that the time to choose is limited?

According to the satisficing model, the agent sets his own aspiration level, in this case, ingredients economically affordable with enough quality to make a dish tasty. The agent then only requires the skill necessary to identify cues in the environment to realize whether an option satisfies his aspirations: is it affordable? Is its quality good enough? According to Simon, the first option that meets the aspiration level is accepted. It is bought, and then, the agent goes for the next ingredient in his list. This agent does not have enough time to look for the optimal choice available in the market for each ingredient of the list comparing each available option for each case. His search is not exhaustive for limitations of time. Also, his cognitive energy es finite. Hence, agents' decision-making does not guarantee the optimal solution for each choice. It is just like Simon (1985, p. 15) asserted, applying the satisficing model to himself: "[...] Since my world-picture is only a crude approximation to reality, I cannot aspire to optimize anything: at most, I can aim at satisficing. Searching for the best can only dissipate scarce cognitive resources. The best is enemy of the good".

In the case above, we see an agent that is neither going to organize all preferences considering price-quality relations, nor is he going to calculate the expected utility of each option to make the decision of highest value. As the choosing time is limited, it is reasonable to accept the idea that agents only care about good enough decisions, not optimization, because it is a waste of energy and time. Satisficing might be understood, then, as a behavior pattern that is very common in daily life situations.

The neoclassical economist would say that the agent always chooses the best option among a limited set of options. According to NE, in some sense, the agent is optimizing when choosing the best option, he has at his hand. But this assumes that agents are utility maximizers who are always looking for the optimization of

their utility functions. This involves the idealization of assuming that agents are fully rational. But this "as-if assumption" keeps the "black box of cognition" closed (Berg & Gigerenzer, 2010), hiding the intrinsic motivations that exist in the behaviors, reasoning and intuition involved in each decision made by agents. We should recall that Simon was interested in modeling an intuition that actually plays a role in decision-making: satisficing, where the intrinsic motivation is just making 'good enough' decisions. Optimization, in contrast, is not a habitual intrinsic motivation in daily life situations. But it might be relevant in a set of well-defined situations as we will see in the fourth section.

Satisficing is a behavioral model that might describe human behaviors better than NE does in some contexts. It was a first step towards a more empirically based economics (Hortal, 2017), where, according to Simon (1985, p. 18), his approach was "[...] abandoning the priorism of classical and neoclassical assumptions about human behavior". This abandoning of the priorism has to do with the idea that assumptions, from an a priori methodological stance, are self-evident statements (axioms in a classical sense) from where inferences are made. And their truth does not depend on experience, but it is completely independent of it. Just like an analytical statement of a triangle that has three sides. The truth is matter of definition. For Simon, in contrasts with this apriorism, all statements and assumptions used in empirical sciences like economics (and artificial intelligence) should be empirically grounded on evidence. Where assumptions about rationality should have empirical support on psychology, artificial intelligence, observational data, and results of experiments that study human behaviors. This is why Simon rejected NE assumptions of rationality to promote a more empirically grounded view of such assumptions that he labeled as bounded rationality.

In this scenario, it is possible to claim that satisficing might substitute NE models in some descriptive contexts. But satisficing is not the only behavioral model available because there is heterogeneity in the practice in BE. Some BE models such as prospect theory (Kahneman & Tversky, 1979), heuristics and biases (Tversky & Kahneman, 1974), limited self-control (Thaler, 1980) explicitly aim to substitute NE at a descriptive level, but not necessarily at a normative level. Other behavioral models are used to describe collective behaviors better than NE does, through concepts such as social norms (Bicchieri, 2017; Sen, 2005; Sunstein, 1996) and narratives (Akerlof & Shiller, 2010; Shiller, 2019). Complementarity links of BE to heterodox approaches have been explored, as the relation between the availability and anchoring heuristics with the Veblenian notion of habits and intuition (Taioka, Almeida & García-Fernandez, 2020). And the same might be done with NE, where optimization might be retained at a normative level, but picking up BE insights to study agents´ descriptive misbehaviors. The main lesson here is the following: we might describe human behavior as utility maximizing, but this it is not the only model available for describing. In BE there is a wide variety of models that describe human behaviors better than NE models do, especially daily-life situations, where, as Laibson and List (2015, p. 386) put it the Principe 1 of BE: "People try to choose the best feasible option, but they sometimes don´t succeed".

Explaining human behavior

In this section, we are focusing our analysis only on micro explanations, because both NE and BE are concern with individual behaviors. As we saw in the second section, explanation involves the exposition of causal relationships, although not all economists explain in this manner as we already said. This is why is relevant to assume heterogeneity in economists' scientific practices where models pursue different epistemic goals. In this context, while it is true that the models can make good predictions (as we will see in the next section), a problem NE models of rational agency face is that such models do not explain causally human behaviors, because they don't show causal relations or mechanisms. I will elaborate on this idea in this section.

As Jon Elster has shown, the "*Homo economicus*" (also named "Econ" by Thaler for short) rests upon two key assumptions: full rationality and self-interest (Elster, 2011). When it is assumed that agents are perfectly rational, in reality there is an idealization (Mäki, 1992) that omits the agent from the analysis in order to focus solely on the structure of incentives and distribution of probabilities of a specific situation. Consequently, what is analyzed in isolation is the assumption of self-interest relative to the different punishment-reward relationships (incentives) of an environment.

In NE models of rationality, there is a methodological assumption that plays a key role in this type of modeling: a reductionist assumption that seeks to reduce all human motivations to mere self-interest. According to Gary Becker, his NE models do not make assumptions about singular motivations, but he has a method of analysis where "The analysis assumes that individuals maximize welfare as they conceive it, whether they be selfish, altruistic, loyal, spiteful, or masochistic. Their behavior is forward-looking, and it is also assumed to be consistent over time. In particular, they try as best they can to anticipate the uncertain consequences of their actions" (Becker, 1993, p. 386). Consequently, utility functions are postulated with different parameters that might incorporate a different class of attitudes, preferences, values, and calculations that lead neoclassical economists to extend their kind of analysis (their rational choice model) to non-traditional economics domains (e.g., sociology, law, political science, and history). This is to achieve the greatest generality in the social sciences through the application of a very simple model of rationality.

The problem with this way of "explaining" is that self-interest and utility functions become notions that are too general to tell us anything about the specific behaviors. As Geoffrey Hodgson correctly said, "[...] the universality of the theory is a sign of weakness rather than strength [...] rational choice theory fails to focus on key human concepts such as culture and learning [...] it fails to identify crucial aspects and become incapacitated by their over-generality" (Hodgson, 2012 pp. 1-2). Conceptual precision is lost, and the risk of misunderstanding certain human behaviors is high. Moreover, when it is explicitly claimed that NE models are not interested in studying agents' motivations to act. To really understand why an agent

makes a specific decision is key to studying axiological rationality, i.e., the values that provide the reasons why agents act as they do (Alvarez, 2016; Bicchieri, 2017; Echeverria & Alvarez, 2008). But these reasons and motivations are left aside in NE models by means of *as if assumptions,* which are not interested in studying the how, i.e., how agents make decisions (Aumann, 2019; Berg & Gigerenzer, 2010). The hiding of cognitive mechanisms here is a consequence of a specific way of modeling rationality, and the human mind is addressed as a black box. It is ironic that rational choice models do not have anything to say about agents' reasoning.

But there are more consequences: neoclassical models tend to absolutize the notion of self-interest (in the form of a methodological assumption as individual utility maximizers") when interpreted as an axiom, a structural assumption that is already given, self-contented, and self-determined (Hodgson, 2012). The notion of self-interest is presented as a self-evident truth applied to analyze the behavior of every rational person, as if it were the main motivation of human action, hiding other motivations that might be relevant to explain human behavior such as emotions (Thaler, 2015) and commitments with social norms (Sen, 2005, 1977). The neoclassical model, with the overgeneralization of self-interest (Becker, 1993; Lazear, 2000), tends to make other motivations invisible, hiding the fact that we are socially interdependent agents (Álvarez, 2016), corporeal and vulnerable bodies because of our cognitive biases (Thaler & Sunstein, 2008).

As Uskali Mäki (2018) would put it, it is very common for economists to misunderstand their models and misuse them. The problem is not in the model's use of false assumptions, but rather in the modeling pragmatics. That is, in the decontextualized uses of models, where it is possible to believe that the same model can be used to achieve any epistemological objective. "Economists excel in manipulating mathematical model descriptions but are much less competent in regard to the details of, and coordination between, the other components of the modelling endeavor – targets, purposes, audiences, issues of relevant resemblance- and this shows in their limited model commentaries." (Mäki, 2018, p. 7).

The great historical merit of BE is that it provides elements that serve to demystify the notion of the rational agent. That is: to de-idealize *Homo economicus* through the incorporation of more realistic or sound assumptions of decision-making, based on psychological grounding (Heidl, 2016). Just in the direction of addressing issues of relevant resemblance, to advance towards the formation of a pattern of causal explanations of agents´ behavior. This is because it studies the causal cognitive processes involved in decision making (Heidl, 2016). And it does so by taking the corporeality of agents and the situated character of rationality seriously, to open the so-called black box of cognition (Berg & Gigerenzer, 2010).

The new notion of the agent of BE, unlike the neoclassical notion, is not based on mere axioms postulated a priori (independent of experience). On the contrary, it is based on robust, plausible psychological theories, far removed from psychological behaviorism (Angner & Loewenstein, 2012), where the cognitive processes of the agent involved in its decision making are taken seriously, i.e., the BE notion of agent is a kind of naturalized abstraction of rational agency.

One of these new abstractions of the rational agent is now synthesized in the dual system of reasoning. This theory postulates that human reasoning is composed of two systems: system 1 and system 2 (Kahneman, 2011). System 1 makes quick inferences. We use it when we touch a hot surface and remove our hands, or when we make quick inferences (like solving the sum 2 + 2). It is automatic, emotional, and its function consumes a few quantities of energy. System 2, on the other hand, is a more deliberative, calculating, and analytical system. It works in a slower manner, and its operation requires too much consumption of energy (Kahneman, 2011). According to BE, most people use system 1 most of the time because it is easier to use. In contrast, system 2 requires a greater input of energy and time (Kahneman, 2011; Thaler, 2015; Thaler & Sunstein, 2008).

This theory raises the possibility of reducing NE to a more general theory of rationality by presenting it as a special sub-case of the dual theory of reasoning. This is suggested by Thaler and Sunstein when they say that:

> [...] To qualify as Econs, people are not required to make perfect forecasts (that would require omniscience), but they are required to make *unbiased forecasts*. That is, the forecasts can be wrong, but *they can´t be systematically wrong in a predictable direction*. Unlike Econs, *Humans predictably err* [...] *Econs respond primarily to incentives* [...] *but they are not influenced by such "irrelevant" factors as the order in which options are displayed. Humans respond to incentives too, but they are also influenced by nudges*" (Thaler & Sunstein, 2008, pp. 7-8) "[...] *Econs never make an important decision without checking with their Reflective Systems* (Thaler & Sunstein, 2008, p. 22; the emphasis is mine).

From the above, we can say that the agents idealized by the economists (the Econs) are agents that only use system 2 ("reflective system"), but not system 1 (Thaler, 2015; Thaler & Sunstein, 2008), because "Econs" make unbiased forecasts. Econs cannot be systematically wrong in their predictions, they must respond only to incentives (i.e., systems of punishment and rewards) and they also have unbounded self-control. The implication is, therefore, that neoclassical agents are not humans that may make systematic mistakes or are influenced by irrelevant factors (as in the displaying of the different options) because of their cognitive biases and the relevance their emotions have in making decisions.

This opens the possibility of reducing NE to system 2 but maintains both systems 1 and 2 as a more general theory of rationality, whereas NE models maintain that agents are perfectly rational, and only uses system 2, while omitting system 1. When system 1 is incorporated in the analysis of agents' behavior, we could then say that there is a process of de-idealization of the rational agent (Heidl, 2016) because intuitions, emotions, heuristic reasoning, and cognitive biases are incorporated into the analysis.

In contrast, NE cannot reduce all the elements that are part of system 1 down to the traditional rational choice model, particularly intuitions, heuristic reasoning,

and cognitive biases. And this is not possible because NE models of rationality are not interested in studying agents' reasoning. Neoclassical economists are only interested in studying the outcomes of a decision-making process, not the *how*, not the manner in which agents make decisions (Aumann, 2019; Gigerenzer, 2015).

The elements given in these subsections allow us to affirm that BE can replace NE in two specific contexts: the descriptive and explanatory levels. This is because it describes the decision-making of real agents (not fictitious agents) better and advances towards a causal explanation of the decision-making of agents, where, in turn, it allows the incorporation of NE into a more general theory of rationality[2]. This incorporation opens the possibility of a strong complementarity of NE and BE in explaining human behaviors in a breadth and deep manner (Marchionni, 2008). A case that might exemplify this kind of complementarity is corruption, as we will see in the next section.

THE EPISTEMOLOGICAL CONTEXTS WHERE BE AND NE MIGHT BE COMPLEMENTARY

Predicting human behavior

It seems that the current behavioral revolution is similar to scientific revolutions that have occurred in other scientific disciplines, such as in physics. When quantum mechanics and general relativity theories arose in the first half of the Twentieth century, neither completely substituted classical mechanics. It is true that Newtonian mechanics was no longer considered a universal and valid theory for all domains in physics, which unified and reduced all movement of all kinds of bodies (atoms, apples, stars) to some "fundamental laws" (Cartwright, 1999; Mitchell, 2009). What happened is that the new disciplines addressed different problems that were unsolvable through the Newtonian theoretical framework. Quantum mechanics, Newtonian physics, and the general theory of relativity are, in fact, situated at different levels of abstractions and answer different problems (Cartwright, 1999; Mitchell, 2009). And up until now there is not a single approach that has been successful in unifying these three fields into a single theory of everything (Cartwright, 1999; Mitchell, 2009). In such a context, scientific revolutions in physics actually lead the physicists to have a more modest view about the scope and limitations of classical mechanics. Physicists had to learn how to think in rather contextual terms (Cartwright, 1999; Mitchell, 2009).

Classical mechanics is not the fundamental or the general theory in physics,

[2] The dual system is not the only general theory of rationality currently available. There are other general theories such as fast and frugal heuristics (Gigerenzer, 2015) and the projects known as 4E where cognition is understood as something embodied, embedded, extended and enacted (Jaegher & Rohde, 2010). Here I used the dual system just to make an argument in favor of the possibility of reducing NE to a more general theory of rationality.

such as Newton used to think, but it is useful for studying some phenomena in some specific domains. Moreover, is useful to develop certain kinds of technology (such as rocket launches, the building of bridges, etc.). Perhaps something analogous is occurring with NE and the behavioral revolution. First, because NE is no longer considered the universal or the most fundamental theory of human rationality. Second, the domains of its applications now are more clearly defined. And finally, third, because after all, the emergence of new behavioral domains and a new set of theories, techniques, and methods do not invalid the development of some theoretical tools that might lead to the development of behavioral technologies. This is the case of the traditional notion of incentives, that are still a powerful tool for intervening in agents' behavior.

In this context, is important to recall a characterization of NE made by Elinor Ostrom:

> [...] Theories are efforts to build understanding by making core assumptions about specific working parts of frequently encountered phenomena and predicting general outcomes. Models are very specific working examples of a theory – and they are frequently confused with being theories themselves. As Alchian [...] pointed out long ago, what is called *"rational choice theory" is not a broad theory of human behavior but rather a useful model to predict behavior in a particular situation – a highly competitive market for private goods* (Ostrom, 2009b, p. 430; the emphasis is mine).

As we can see, for Ostrom, NE models are not a broad or general theory of human behavior, but rather a useful model that might predict behavior in a specific situation: a highly competitive market for private goods. This idea is important because it shows us that NE models are not so general and fundamental as neoclassical economists used to think, just as was the case in physics with classical mechanics. The flexible character of the neoclassical model has found a niche, where it has become entrenched, since this model is useful for predicting individualistic behavior in competitive contexts (Ostrom, 2009a). But for understanding and forecasting cooperative and pro-social interactions, perhaps we should rely on a more general theory of rationality, not one that is so partial and unilateral, as is the case with NE.

It seems that BE might be more general because it is possible to reduce NE to the dual system of reasoning as we saw in the third section. But in spite of this, NE might still work well in some limited contexts. And these contexts are linked to the epistemological objective of prediction. Neoclassical economists recognize that real agents are not perfectly rational, but they assume that they act as if they were. Thus, the notion of a neoclassical agent is a mere instrumental assumption that serves to formulate predictive mathematical models. But this predictive character, as I showed in the third section, is weaker than the neoclassical economists tend to claim. That is why it is better to talk about forecasts (Hausman, 1992; Lawson, 1997).

Neoclassical models of rationality tend to make good forecasts on competitive

and individualistic contexts. A case in point is the subject of corruption. For example, if we have traditional politicians who are already corrupt and situated in competitive contexts (where they compete with other corrupt politicians for economic resources), we can say that if the incentive structure of a specific situation makes it profitable to violate legal and moral standards, and there is little likelihood of being caught and punished, there will be more politicians who will act corruptly (Rose-Ackerman, 1975). This forecast has some empirical support if we consider contexts such as the lack of transparency, inefficiency in justice systems, and impunity in some countries (Rose-Ackerman, 1975).

At this point, it is important to note, however, that these models are not describing the behavior of real agents but rather telling us what certain agents can do, given a specific socio-institutional context. The focus of the analysis, again, is not the agent, but rather the incentive structure of a specific socio-institutional context. One of the problems is that neoclassical models assume that we are all potential rogues (Ariely, 2008, 2012); that everyone can be a transgressor of norms (legal and moral), as long as the transgression is profitable (Becker, 1993; Rose-Ackerman, 1975). While it is true that harder punishments might diminish criminal and corrupt behavior, this is not the whole story.

BE has also worked on the issue of corruption and dishonesty by showing, without denying what has been stated by neoclassical economists, that there are other factors that explain corrupt behavior: corrupt slippery slopes, a gradual change of language, and behavior based on imitation, among other aspects (Ariely, 2008, 2012). It has also been documented, based on experiments, that agents may manifest dishonest behavior if presented with the opportunity, but that they do not manifest the dishonest behavior to the degree predicted by the neoclassical theory (Ariely, 2008, 2012).

BE does not deny the successful forecasts of NE made under specific situations. But since there can be a reduction of the neoclassical model to the dual system of the mind, it can be said that there is a complementarity. There is nothing in the dual system that denies the possibility of success of NE in certain specific contexts. On the subject of corruption, as we can see, there is no competition, but rather a complementarity between BE and NE because both of them are situated at different levels of abstraction. The NE approach to corruption studies the punishment-reward systems (incentives) that are part of a social context (NE), omitting the agent, while BE studies the corrupt decision making (or corrupt reasoning) taking into account psychological insights (like imitation, cognitive biases, heuristic reasoning, the lack of self-control) that help us understand corrupt behaviors. BE studies a kind of corrupt reasoning while NE studies a kind of optimal corrupt behavior defined by incentive structures. NE seeks to change the perverse incentives that lead to a reproduction of corrupt behaviors, while BE seeks to change the reasoning of potential or already corrupt agents.

We can see a complement between BE and NE. But again, we should emphasize modesty: NE models usually work well, but they are fallible and partial epistemic devices. Not an infallible general theory of rationality. In this sense, NE might

predict a bounded range of behaviors. Not all behaviors, especially those related to system 1 of reasoning where lack of self-control, heuristics, cognitive biases, social preferences, and emotions are situated. In this cognitive area, BE models might be used. But also, modesty should be emphasized for behavioral models: There are some domains where NE still works well, especially in those behaviors related to system 2 of reasoning situated in competitive environments, where traditional norms of rationality are still relevant in some contexts. However, much work about systematization has to be done to achieve a systematic general theory of rationality, where it should be shown how NE and BE might be complementary recognizing plurality in economists' scientific practices.

Prescribing optimization

Although NE does not describe and explain real human behavior, it does discuss how agents should behave, especially under competitive contexts. For example, if we are in a strategic interaction and we want to win a game, minimize costs, or simply avoid defeat, we can apply traditional game theory to evaluate scenarios and make decisions about our strategies, considering the best strategies that other agents may apply. This does not mean that real people always make rational strategic decisions, but "[...] a rational player must take into account how other players reason about him" (Aumann, 1987, p. 17). So, in such competitive strategic interactions, optimization plays a normative role which consists of not being overconfident about our strategies and not underestimating the skills of other agents.

There are other contexts where being consistent with game theory (and rational choice models) would lead agents to suffer terrible losses. For example, being individualistic in a cooperative context can lead agents to maximize their individual utility in the short term, but if other agents discover it, they will apply actions against free riders. This is because cooperative agents reject agents who act based solely on self-interest. Reciprocity here is important because when an agent is not pro-social in a cooperative context, he will be punished and excluded from the pool resources. Elinor Ostrom has shown several examples where local social systems have their own institutions (rules of interactions) that self-regulate the system (Ostrom, 2009a, 2009b), sometimes with better social outcomes than market mechanisms or public actions performed by the governments.

The role of social norms (or institutions) is stronger than neoclassical economists often believe because they have a deontic component and also enable sociability (Bicchieri, 2017). Social norms provide information about what to do and what not to do in certain contexts, and also serve to anticipate the consequences of agents´ acts: if an act is performed in a specific context, then a response is expected from others (Bicchieri, 2017). Social norms matter, in such a way that is not always the best option for an agent to try to maximize his individual self-interest omitting these norms. Sometimes agents should and ought to behave following through on their commitments in accordance with social and moral rules (Sen, 2005).

In other contexts, optimization is not only desirable, but it is what must be done to achieve a goal. One example is the role linear programming has in the

process of planning. According to Leonid Kantorovich, "[...] The optimizing approach is here a matter of prime importance. The treatment of the economy as a single system, *to be controlled toward a consistent goal, allowed the efficient systematization of enormous information material, its deep analysis for valid decision-making.*"[3] (the emphasis is mine). Here we can see that optimization is important for normative reasons: for consistently achieving goals and for making valid decisions. But, as was the case with Kantorovich, in some circumstances, optimization is what must be done to survive. During the siege of Leningrad (now Saint Petersburg) in World War II, Soviet authorities called Kantorovich to calculate the temperature and thickness of ice that would support trucks of a certain weight and speed passing over a frozen lake, in order to bring supplies to the Soviet city (Bollard, 2019). Kantorovich had to calculate the optimal distance between a number of Soviet tanks, consider changes in environmental conditions, etc. He developed an algorithm that helped the soviets use the "road of life" (as the frozen lake was called) to bring supplies and, later, evacuate citizens. "He was frequently to be seen out on the lake testing the conditions, sometimes walking between the vehicles". (Bollard, 2019, p. 138). In this dramatic situation, we can see that optimization was a matter of life or death.

But optimization is not always what must be done, sometimes because it is simply impossible to optimize. Regarding this point, we should recall the distinction between well and ill-formulated problems elaborated by Simon & Newell (1958):

> In short, *well-structured problems are those that can be formulated explicitly and quantitatively, and that can then be solved by known and feasible computational techniques*. What, then, are ill-structured problems? Problems are ill-structured when they are not well-structured. In some cases, for example, *the essential variables are not numerical at all, but symbolic or verbal*. An executive who is drafting a sick-leave policy is searching for words, not numbers. Second, *there are many important situations in everyday life where the objective function, the goal, is vague and nonquantitative*. How, for example, do we evaluate the quality of an educational system or the effectiveness of a public relations department?' Third, there are many practical problems it would be accurate to say 'most practical problems'- for *which computational algorithms simply are not available. If we face the facts of organizational life, we are forced to admit that the majority of decisions that executives face every day and certainly a majority of the very most important decisions lie much closer to the ill-structured than to the well-structured end of the spectrum* (Simon & Newell, 1958, p. 5; the emphasis is mine).

As we can see, the "well-structured problems" are those to which it is possible to give a mathematical form because its structure follows linear tendencies that

[3] See Kantorovich (1992).

could be formulated explicitly and quantitatively. Such problems are solved by optimization techniques, such as linear programming, research operation, algorithms, etc. But there are "ill-structured problems" that are not possible to formulate in mathematical terms because their variables are verbal and symbolic, the goal is vague and non-quantitative, and computational algorithms are not available. Heuristic reasoning, as opposed to algorithms, is required here for solving problems. Heuristics may help solve a problem, but they are fallible rules of thumb that involve the use of intuition, insight (or perception), and learning (Simon & Newell, 1958). Their use might be normative to solve ill-structured problems, especially when the time to make a decision that solves the problem is limited and the agent is under pressure by the environment. Perhaps the use of heuristics does not provide an optimal, unique, and infallible solution, but they may be *good enough*, as Simon used to say.

As we have seen in this section, optimization and utility-maximizing behaviors are normatively valid only for some limited contexts, where agents face well-structured problems (such as those solved by Kantorovich and his linear programming techniques) and competitive situations where agents are under strategic interactions (such as those addressed by Aumann in his notion of correlated equilibrium). But optimization techniques are not valid for all contexts, such as with ill-structured problems (addressed by Simon and the followers of bounded rationality) and with cooperative situations where agents interact under social norms (addressed by Ostrom and Sen). In such situations, we need other theoretical tools such as heuristics (which might be normative) and the incorporation of pro-social preferences.

FINAL REMARKS

In this article, we have addressed the problem of substitution-complementarity between NE and BE in the context of the current behavioral revolution. To do so, we first distinguished different epistemological contexts involved in the research practices of economists: description, explanation, forecasting, and prescription of what agents should do and what policy advises. We showed that the substitution-complementarity issue might have varying impacts and expressions in these different contexts. From these elements, we arrive at the following conclusions.

The first conclusion is that any attempt of "synthesis" between NE and BE cannot be harmonious because economics is a very complex science that involves several epistemic contexts and levels of analysis. Second, it seems that there are elements to assert that NE might be reduced as a special case of a rather general theory of rationality developed by BE, where at a descriptive and explanatory level, BE could substitute the NE, but at a predictive and prescriptive level, there may be complementarity because there are some domains where NE still works well. In this sense, NE might complement BE. Finally, further research should be done to show the impact of BE on other subdisciplines of economics. In the very least the complexity of human behavior is increasingly recognized and new behavioral tools are at the disposal of social scientists to understand and for changing behaviors.

REFERENCES

Akerlof, G. A., & Shiller, R. J. (2010). *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism*. Princeton University Press.

Alvarez, J. F. (2016). Conflicts, bounded rationality and collective wisdom in a networked society. In *Paradoxes of conflicts* (pp. 85–95).

Angner, E. (2019). We're All Behavioral Economists Now. *Journal of Economic Methodology, 26*(3), 16.

Angner, E., & Loewenstein, G. (2012). Behavioral economics. In *Philosophy of Economics*.

Ariely, D. (2008). *Predictably irrational: The hidden forces that shape our decisions* (1st ed). Harper.

Ariely, D. (2012). *The (honest) truth about dishonesty*. New York NY: Harper Collins Publishers.

Aumann, R. J. (1987). Correlated Equilibrium as an Expression of Bayesian Rationality. *Econometrica*, *55*(1), 1–18. JSTOR. https://doi.org/10.2307/1911154

Aumann, R. J. (2019). A synthesis of behavioural and mainstream economics. *Nature Human Behaviour, 3*(7), 666–670. https://doi.org/10.1038/s41562-019-0617-3

Becker, G. (1993). Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy, 101*(3), 385–409.

Berg, N., & Gigerenzer, G. (2010). As-if behavioral economics: Neoclassical economics in disguise? *History of Economic Ideas, 18*(1), 133–165. JSTOR.

Bicchieri, C. (2017). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.

Bollard, A. (2019). *Economists at War: How a Handful of Economists Helped Win and Lose the World Wars*. Oxford University Press.

Bowles, S. (2014). Niccolò Machiavelli and the Origins of Mechanism Design. *Journal of Economic Issues, 48*(2), 267–278. https://doi.org/10.2753/JEI0021-3624480202

Bowles, S., & Polania-Reyes, S. (2012). Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature, 50*(2), 368–425.

Chetty, R. (2015). Behavioral economics and public policy: A pragmatic perspective. *American Economic Review, 105*(5), 1–33.

Daniel Kahneman. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *The American Economic Review, 93*(5), 1449–1475.

Echeverria, J., & Alvarez, J. F. (2008). Bounded rationality in social sciences. In *Epistemology and the social* (pp. 173–191).

Elster, J. (2011). *El desinterés: Tratado crítico del hombre económico (I)*. Siglo XXI.

Gigerenzer, G. (2015). *Risk savvy: How to make good decisions*. Penguin.

Grant, R. W. (2002). The ethics of incentives: Historical origins and contemporary understandings. *Economics & Philosophy, 18*(1), 111–139. https://doi.org/10.1017/S0266267102001104

Hacking, I. (1983). *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge University Press.

Hausman, D. (1992). *The inexact and separate science of economics*. Cambridge University Press.

Heidl, S. (2016). *Philosophical Problems of Behavioural Economics* (1st ed.). Routledge. https://doi.org/10.4324/9781315637372

Hodgson, G. M. (2012). *From pleasure machines to moral communities: an evolutionary economics without homo economicus*. University of Chicago Press.

Hortal, A. (2017). Empiricism in Herbert Simon: Administrative Behavior within the evolution of the Models of Bounded and Procedural Rationality. *Brazilian Journal of Political Economy, 37*(4), 719–733. https://doi.org/10.1590/0101-31572017v37n04a04

Hortal, A. (2019). Nudging and educating: Bounded axiological rationality in behavioral insights. *Behavioural Public Policy*, 1–24.

Jaegher, H. D., & Rohde, M. (2010). *Enaction: Toward a New Paradigm for Cognitive Science*. MIT Press.

Jolls, C., Thaler, R., & Sunstein, C. R. (1998). A Behavioral Approach to Law and Economics. *Stanford Law Review, 50*, 80.

Kahneman, D. (2011). *Thinking fast and slow*. Macmillan.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*(2), 263. https://doi.org/10.2307/1914185
Kantorovich, L. (1992). Mathematics in Economics: Achievements, difficulties, perspectives. In *Nobel Lectures, Economics 1969-1980*. World Scientific Publishing Co.
Lawson, T. (1997). *Economics and reality*. Psychology Press.
Lazear, E. (2000). Economic imperialism. *The Quarterly Journal of Economics, 115*(1), 99–146.
Laibson, D., & List, J. A. (2015). Principles of (behavioral) economics. *American Economic Review*, *105*(*5*), 385-90.
Mäki, U. (2018). Rights and wrongs of economic modelling: refining Rodrik. *Journal of Economic Methodology*, *25*(3), 218-236.
Mäki, U. (1992). On the method of isolation in economics. *Poznan Studies in the Philosophy of the Sciences and the Humanities, 26*.
Marchionni, C. (2008). Explanatory pluralism and complementarity: From autonomy to integration. *Philosophy of the Social Sciences*, *38*(3), 314-333.
Mitchell, S. (2009). *Unsimple truths: Science, complexity, and policy*. University of Chicago Press.
Morgan, M. S. (2014). What if? Models, fact and fiction in economics. *Journal of the British Academy, 2*, 231–268.
Ostrom, E. (2009a). *Nobel prize lecture*.
Ostrom, E. (2009b). A general framework for analyzing sustainability of social-ecological systems. *Science*, *325*(*5939*), 419-422.
Reiss, J. (2013). *Philosophy of Economics. A contemporary introduction*. Routledge Taylor & Francis Group.
Rose-Ackerman, S. (1975). The economics of corruption. *Journal of Public Economics, 4*(2), 187–203.
Sen, A. (2005). Why exactly is commitment important for rationality? *Economics and Philosophy, 21*(1), 5–14. https://doi.org/10.1017/S0266267104000355
Sen, A. K. (1977). Rational Fools: A Critique of the Behavioral Foundations of Economic Theory. *Philosophy & Public Affairs, 6*(4), 317–344. JSTOR.
Shiller, R. J. (2019). *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*. Princeton University Press.
Simon, H. A. (1985). My life philosophy. *The American Economist*, *29*(1), 15-20.
Simon, H.A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*(2), 129–138. https://doi.org/10.1037/h0042769
Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69(1), 99-118*.
Simon, H. A., & Newell, A. (1958). Heuristic problem solving: The next advance in operation research. *Operation Research*, *6*(1), 1–10.
Sunstein, C. R. (1996). Social norms and social roles. *Columbia Law Review, 96*(4), 903–968.
Taioka, T., Almeida, F., & Fernández, R. G. (2020). Thorstein Veblen's Institutional Economics and Daniel Kahneman and Amos Tversky's Behavioral Economics: an analysis of convergent points. *Brazilian Journal of Political Economy*, *40*(4), 746-765.
Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*(1), 39–60.
Thaler, R. (2015). *Misbehaving: The making of Behavioral Economics*. Penguin.
Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124