

Department of Economics
Working Paper

**Belief Reversals as Phase Transitions and Economic
Fragility:**

**A Complexity Theory of Financial Cycles with
Reflexive Agents**

By

John B. Davis

Working Paper 2020-02

College of Business Administration



Forthcoming in *Review of Evolutionary Political Economy*

Belief reversals as phase transitions and economic fragility:

A complexity theory of financial cycles with reflexive agents

John Davis, Marquette University and University of Amsterdam

“By the way I never did grasp just what reflexivity was” (Samuelson, 1993, p. 7).

Abstract: This paper aims to contribute to the analysis of expectations and belief reversals in an evolutionary and complexity economics framework. It formulates its analysis in terms of the concept of reflexivity, drawing on the ideas regarding reflexivity in financial markets of George Soros, and lays out a model of how a financial cycle expresses a systematic pattern of interacting feedback effects. The paper develops this analysis as a complex interaction between sets of heterogeneous expectations derived from the behavior of reflexive economic agents. Positive and negative feedback phases in a cycle are distinguished and associated with boom and bust stages of that cycle. A central role is played by agents’ beliefs and judgments underlying their expectations, and how those beliefs and judgments in uncertain circumstances are changeable and subject to abrupt reversals which can manifest themselves in ‘Minsky moments.’ The paper argues that agents’ belief reversals result follow their misconception about causal processes in booms and upswings, a misconception that reflects their tendency to think causally in terms of negative feedback patterns rather than positive ones.

Keywords: reflexivity, feedback, complexity, boom-bust, Soros, judgment, Minsky

1. From Arthur to Keynes to Soros

The current economic crisis, precipitated by the corona virus pandemic, demonstrates the inherent fragility of the contemporary capitalist economic system and the drastic consequences and costs in human suffering that result from neglecting its needed reform. From a period of seeming high employment and feverish speculative activity in financial markets to depression conditions and widespread market crashes, this current economic system is marked by severe and abrupt swings and reversals in expectations and beliefs about so-called economic fundamentals. Yet there is little theory in economics explaining expectation and belief reversals despite the fact that the concept of expectations has been central to thinking about the economy since J. M. Keynes. This paper, then, seeks to contribute to the theory of expectations and belief reversals by analyzing it an evolutionary and complexity economics framework. It formulates its analysis specifically in terms of the concept of reflexivity, drawing on the pioneering ideas regarding reflexivity of George Soros regarding reflexivity in financial markets. Consider first, then, the connection between an evolutionary-complexity approach to the economy and the reflexivity concept.

Brian Arthur characterizes complexity economics as a framework for economic thought fundamentally different from standard neoclassical equilibrium theory in the following way:

It is a different way of seeing the economy. It gives a different view, one where actions and strategies constantly evolve, where time becomes important, where structures constantly form and re-form, where phenomena appear that are not visible to standard equilibrium analysis, and where a meso-layer between the micro and the macro becomes important (Arthur, 2013, p. 1).

Arthur, then, emphasizes the reflexivity connection. In complex economic systems, economic agents' actions and strategies constantly evolve because agents interact in networks of heterogeneous agents who all constantly revise and adapt their actions and strategies in response to how others revise and adapt their actions and strategies. That is, economic agents are adaptive economic agents. What emphasizing reflexivity may explain, then, is how people revise and

adapt their actions and strategies in interaction with one another in continually changing circumstances.

This understanding clearly requires that we make time central to the explanation of the economy. Reflexivity accordingly explains this time dimension in terms of the idea of a temporal feedback loop, whereby people's expectations of their future interaction with others feed back on and influence their current actions and strategies, while their current actions and strategies feed forward on and influence their future interaction with others. In effect, reflexivity makes expectations interdependent across both people and time. We can summarize this by saying that reflexivity theory is a theory of the agency of adaptive agents embedded in evolving structures and networks of social and economic interaction. Following Arthur's comparison of complexity theory and standard equilibrium theory, then, what reflexivity theory accordingly provides us is a relational account of socially and economically embedded adaptive agents compared to neoclassicism's fixed preference account of atomistic agents. I thus characterize these adaptive agents as 'reflexive economic agents' and describe their interaction in this paper in terms of reflexive economic processes.

This paper begins to lay out the role reflexivity theory understood in this way plays in explaining how reflexive economic agents interact in complex economic processes. Where does one start? One of the most important subjects of application of reflexivity thinking in economics is interdependent expectations in financial markets. I begin with this subject – or actually one paradigmatic aspect of it, boom-bust cycles in financial markets – because it allows us to first focus on whole sets of heterogeneous expectations and explain reflexive behavior in a systemic way in terms of the evolution of a specific structure of interaction. A complexity-reflexivity approach nonetheless seeks to explain how structures of interaction and economic agents themselves co-evolve. My procedure, then, is to begin at the level of a structure of interaction to bring out what a reflexive process involves, and then turn to how this influences agents' evolution, before turning back to how this influences the evolution of the structure in which they interact.

Keynes initiated this approach in terms of interdependent expectations when he framed his beauty contest idea in reflexivity terms as a single set of interdependent expectations resulting from investors focusing only on what others believe others believe are good investments

(Keynes, 1936).¹ More recently, Soros, who explicitly characterizes his analysis of boom-bust cycles and bubbles in financial markets as an application of reflexivity theory, links speculative expectations and expectations about economic fundamentals, and uses this to develop his view of how expectations operate in financial markets (Soros, 1987, 2013). It seems, then, that if we are to explain reflexivity in financial markets in terms of interdependent expectations, whatever particular application we choose to adopt, what we first need to do is determine how we should describe different ‘sets’ of expectations. To do this, I begin with Soros’ boom-bust analysis, since in combining speculative expectations and expectations about fundamentals, he provides a comprehensive view of what a set of interdependent expectations might involve. Boom-bust cycles and bubbles, of course, are just one financial markets phenomenon. But for complexity-reflexivity theory they are a paradigmatic one, because they concern how a specific, relatively well-defined structure evolves together with the evolution of agents’ expectations.

For Soros, then, every boom-bust process has two components: an underlying trend in fundamentals and a misconception about value of a financial asset or stock based upon those fundamentals. In the boom, the trend in the underlying value is reinforced by the misconception about the future value of the stock that is driving up its price. The rising stock price improves the fundamentals on which that stock’s value is based, for example, in the form of greater capacity to finance operations or to acquire other firms, driving up earnings per share and thus giving further impetus to the stock’s price rise (see Soros, 2014, p. 324 and Figure 4). In reflexivity terms, speculators’ expectations of the future, in this case involving a misconception about the future value of the asset, feed back on and influences the present by raising the current market price and earnings per share, which then feeds forward on speculators’ expectation of the future, etc., in a continuing reflexive feedback loop.

Yet speculative behavior, Soros argues, drives the stock price up faster than firms are able to improve earnings, and at some point, the price is seen to be too high relative to fundamentals, and the bust occurs. He associates the boom with a positive feedback process in expectations about the stock’s price – its increases lead speculators to expect even higher prices for the stock. After the bust, a negative feedback process prevails, and the stock price increases at a decreasing

¹ Keynes thought more in complexity terms than most researchers recognize (cf. Marchionatti, 2010; Davis, 2017), perhaps because he employed an earlier vocabulary with other associations.

rate, reflecting underlying trends in earnings and the gradual improvement in the stock's fundamentals. These latter circumstances involve what he calls 'near-equilibrium conditions' compared to the bubble-like 'far-from equilibrium conditions' that exist in the boom. In effect, the circumstances of a positive feedback cycle depart from and then collapse back on the circumstances of negative feedback, and the speculative episode of the bubble is replaced by everyday course of events in which businesses operate reasonably predictably.

Soros' focus, then, like Keynes's, is on financial markets and stock prices, but in contrast to Keynes's beauty contest idea, Soros' reflexivity theory makes the boom-bust cycle depend on interrelated movements in stock prices and their underlying fundamentals. Since no one denies, including Keynes (cf. Marcuzzo, 2020), that fundamentals play some sort of role in determining stock prices, his financial market conception of a set of expectations is perhaps better seen as a limiting case.² At the same time, since the determinants of and movements in stock prices and their underlying fundamentals are relatively independent, Soros' analysis suggests that we should think not just in terms of one single set of expectations governing stock prices, but in terms of two different or heterogeneous types of expectation processes, one associated with the financial side of the economy itself, and one associated with economic fundamentals, whose interaction then explains boom-bust cycles, bubbles, and other similar financial phenomena. This perspective makes reflexivity theory a theory of interacting, interdependent expectations, involving a complex interaction of multiple, heterogeneous sets of expectations and different interacting, reflexive economic agents.

I adopt this conception of expectations in this paper and develop it in the most basic way in terms of just the two sets of expectations involved in Soros' analysis of boom-bust cycles. The immediate task of the paper, then, is to make two arguments to explain how these boom-bust cycles work:

² There is of course considerable debate over what fundamentals are, and in the view many critics of the mainstream share their nature is misrepresented in the neoclassical individualist welfare framework. Thus, here I treat the idea of fundamentals very broadly as the main determinants of economic activity in production and consumption, and assume that they evolve together with economic activity.

- first, one about how the two sets of expectations regarding stock prices and fundamentals interact at the turning point in the cycle;
- second, one about how these two sets of expectations interact in the upswing of a boom leading up to that turning point.

The first argument has the goal of showing that the turning point in a boom-bust cycle involves a phase transition or a kind of qualitative change – an important object of explanation in complexity economics – due to a change in the relationship between the two sets of expectations that manifests itself in a wave of what I will call ‘belief reversals.’ Soros doesn’t explain why a bust occurs or why economic fundamentals re-assert themselves following a boom. He only says that at some point agents’ views about stock prices become so far removed from the underlying trends determining stock prices that the bust ensues. But he also implies that the problem is that stock prices are ‘out of balance’ with those underlying trends in fundamentals during the boom. The first part of my argument, then, explains this idea in terms of how these two types of expectations, that have gotten increasingly ‘out of balance,’ get rebalanced at the turning point of the cycle. The importance of explaining turning points in boom-bust cycles in terms of how different sets of expectations get rebalanced for complexity theory, then, is that it identifies one mechanism by which the interdependence of agents produces changes in the structure of the economy, or at least an aspect of it.

The goal of the second argument is to examine why these two types of expectations get ‘out of balance’ in the boom. I follow Soros in explaining booms in terms of misconceptions about the nature of the positive feedback processes operating on stock prices and economic fundamentals but argue that the problem especially concerns positive feedback processes associated with economic fundamentals due to there being an inherent ambiguity involved in economic agents understanding their underlying cause-and-effect determinants. The general problem of understanding the underlying causal determinants of observable processes in the case of negative feedback is much less serious, because these prevail before and after the boom in ‘near-equilibrium conditions’ when economic relationships are relatively stable. I thus argue that this biases economic agents in his ‘far-from equilibrium conditions’ toward seeing positive feedback

processes also as stable and predictable on the model of negative ones, though often they are not, and it is this that explains misconceptions in the boom about stock prices.

What is particularly different about the way in which I approach reflexivity in this paper in connection with financial markets is that I treat expectation as a judgment, or rather as a reflexive judgment. There are two main reasons I take this approach. First, focusing on the judgments people make about the world makes it possible to talk about their beliefs, because people's judgments reflect their beliefs. People's beliefs about the economy also continually change as they interact with one another, and significant change in their beliefs is often associated with significant change in the economy. Thus, focusing on judgment takes us beyond simply mapping out changing sets of expectations, and gives us a way of understanding what motivates change in people's expectations. Second, what is especially important about an act of judgment is its point-in-time nature. A judgment is a singular event or episode in historical time when agents consult their beliefs and take a position in interaction with others in ways that may influence how their interaction plays out. Thus a focus on judgment as a singular occasion on which agents act on their beliefs and possibly change them – a 'Minsky moment' (1982) – makes it possible to investigate both turning points as a unique occasions involving significant shift in beliefs, and the speculative positions they adopt when the evidence for doing so may fall short of justifying those positions.

Focusing on judgment admittedly gives my approach a somewhat different way of explaining complex adaptive economic systems compared to the way in which they are mathematically explained in much of complexity economics in terms of the nonlinear dynamics of agent interaction in networks. For both methods, an important explanatory goal is to explain phase transitions as qualitative changes in the structure of the economy. Phase transitions can be described quite well in mathematical terms, but this still leaves unexplained why a specific *state* of expectations produces a transition to a new state of affairs in the economy. Thus, an "uncomfortable thing" about many such models and explanations "is that there is no necessary proximate major cause for a sudden shift in the aggregate state" that would explain why an economy self-organizes itself on some new basis (Kirman, 2010, p. 502). Thus, one goal of the judgment-based approach to reflexivity employed here is to explain in terms of configurations of beliefs, which we can discuss and evaluate in light of their specific historical circumstances.

Section two lays the basic foundation for the rest of the paper by defining and explaining the nature of a reflexive process with feedback loops, and by emphasizing the unique role that reflexivity plays in social science. Section three, concerning my first argument, describes the turning point in a boom-bust cycle as a phase transition brought about by a change in the relationship between two sets of expectations. I associate this change with a wave of belief reversals and describe how it can be precipitated by self-fulfilling or self-defeating prophecies understood as a special type of reflexive prediction. In section four, concerning my second argument, I argue that the misconception Soros believes operates in the upswing of a boom-bust cycle is due to an inherent ambiguity associated with the nature of positive feedback processes that drive the boom. I place positive feedback processes on a spectrum from relatively predictable to highly unpredictable, and argue that economic agents fail to recognize situations at the latter end of the spectrum because they are biased toward understanding cause-and-effect processes in terms of relatively predictable negative feedback processes. In section five I draw some general conclusions about what reflexivity contributes to our thinking about complex economic processes and about what this implies about the nature of reflexive economic agents.

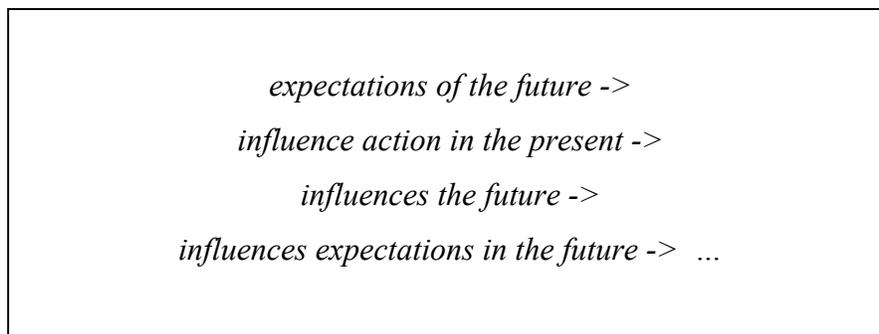
2. Reflexivity in general terms

The basic idea behind the concept of reflexivity is, as the word itself suggests, that of something ‘turned back upon itself.’ In social science the concept refers to how the views or expectations that people have of the future ‘turn back on’ the present and influence what they do in the present. In philosophy of science terms, reflexivity thus complicates our basic one-directional understanding of cause-and-effect in which what happens in the present simply causes what happens in the future. This basic understanding fits the natural world and physical science fairly well, because natural agents have little or no capacity at all to imagine the future, so that views of the future do not feed back onto and influence their behavior in the present. In the social world, however, people can imagine the future, and furthermore can use their understanding of cause-and-effect to change what they do in the present in order to influence what happens in the future. Human beings differ from other natural agents in that they are intentional beings and have a rudimentary understanding of the nature of causality. Because they can anticipate the future, they can form plans for the future, and then reflexively act to promote those plans in the present.

Thus, not only does the present affect the future in the human world, but the future, or rather the views people have of it, ‘turns back on’ and affects the present.

To fill this picture out a bit more in terms of the idea of a feedback loop, we should add that since what people do in the present in anticipation of the future has causal effects on how the future plays out, how the future is changed then enters into how people subsequently view the future. That is, when people act in the present with a view to changing the future, this constantly changes not only how the world works but also ultimately their views of the future as well. In effect, human intentionality and understanding of how the world works imply that the present and the future are both continually being reconstituted through human action. The social world is thus never just as it is but is always emergent and evolving as Arthur emphasizes, or is a non-ergodic process as Paul Davidson describes it (Davidson, 2009). I thus characterize a reflexive process in the social world as an on-going, constantly evolving two-directional cause-and-effect process between the present and the future in which views of the future influence the present, which influence the future, which influence views in the future, and so on. This is set out schematically as a feedback loop in Figure 1.

Figure 1: The reflexive adjustment process



Compared to the more predictable way, then, in which in the natural world what happens in the present basically brings about what happens in the future, in the social world people’s views or expectations of the future are open-ended, and act like a ‘wild card’ in determining how causal

processes play out. This means that there are potentially many different ways in which people might form expectations about the future.³

Nonetheless, to simplify I suggest that in order to explain expectations across boom-bust cycles we differentiate the two types of expectations distinguished above in terms of the two time dimensions in which economic values are customarily recorded, namely, those made in *periodic, point-in-time* terms with regard to stock-type economic values, such as for assets and stocks, and those made in *through-time, rates of change* terms for flow-type economic values, such as are involved in economic fundamentals. This temporal distinction that economics uses surely reflects ways in which people understand time. It also reflects the two sides of the boom-bust cycle that Soros explains in terms of the interrelated movements of stock prices and their underlying fundamentals. Reflexivity theory applied in this way, then, provides a way of analyzing economic agents' different types of expectations according to how those expectations correspond to the main time dimensions of economic values.

3. First argument: Turning points in boom-bust cycles as widespread belief reversals

Here I explain how expectations about stock prices and expectations about fundamentals change in relation to one another at turning points in boom-bust cycles when they are the most 'out of balance' in order to explain turning points as phase transitions as widespread belief reversals. I begin my discussion of the boom-bust cycle with its phase transition turning point because phase transitions often reveal undetected structural features of systems not visible at other points in their dynamics. The undetected structural feature in this case is the reversal of order of judgments about stock prices and fundamentals at the turning point that ends the boom and inaugurates the bust. In the discussion in the next section on the boom, then, I in effect work backwards, to analyze how expectations about stock prices and fundamentals could have gotten into their prior state if they were to subsequently reverse and unravel as they do at the turning point.

³ In effect, they engage in counterfactual rather than Bayesian reasoning what the facts are is always being reconsidered (cf. Davis, 2018).

Soros' boom-bust cycle, we saw, has three phases: the 'far-from-equilibrium' upswing, the phase transition turning point, and the downswing and return to 'near-equilibrium' conditions. The turning point – the 'Minsky moment' – is clearly connected to an unexpected change in agents' beliefs. Arthur argues that we live "in a world where belief, strategies, and actions of agents are being 'tested' for survival within a situation or outcome or 'ecology' that these beliefs, strategies and actions together create" (Arthur, 2013, p. 5). The 'test' of beliefs that I examine occurs when a judgment and prediction about the future in the form of a self-fulfilling or self-defeating prophecy is made that causes agents to reverse their beliefs and produces the turning point in the cycle. Then agents' two types of expectations, which had grown increasingly 'out of balance' in the 'far-from-equilibrium' upswing, unexpectedly get rebalanced toward one another, leading the cycle into the downswing.

There is a long history of research in reflexivity theory on self-fulfilling and self-defeating prophecies as two related forms of a single kind of reflexive judgment that has the potential to produce belief shifts. Self-fulfilling and self-defeating prophecies are defined as a special type of prediction that once expressed and known to others cause people to change their beliefs and behavior in such a way as to either fulfill the prediction, when what it predicted would otherwise not have occurred, or overturn and defeat the prediction, when what it predicted would otherwise have occurred. They are different from the more ordinary sort of predictions which do not significantly alter people's beliefs or alter the course of events, and which dominate much of the economic landscape and conventional analysis of expectations in 'near-equilibrium conditions.'

From a reflexivity perspective, self-fulfilling and self-defeating prophecies are especially interesting because they provide a means of making phase transitions central to the analysis of capitalist market economies. Too long capitalist market economies have been described as evolving in a steady, gradual manner rather than as a system inherently prone to unexpected disruptions that have costly effects on people's well-being. In connection with financial markets, whose central role in capitalist systems is beginning to be better appreciated, self-fulfilling and self-defeating prophecies points us to the delicate balance between institutional structures and agents' beliefs. I suggest, then, that their examination provides a useful means of investigating a range of other potentially fragile structures in capitalist market economies.

i. Self-fulfilling prophecies

The self-fulfilling prophecy idea was used by the sociologist Robert K. Merton to explain how bank runs occurred in the Depression in the U.S. in the 1930s (Merton, 1949).⁴ A self-fulfilling prophecy involves a judgment and a prediction that sets in motion a reflexive process whereby a false belief expressed in the prediction is transformed into a true belief through the influence that it has on others. In Merton's hypothetical example, a bank analyst reviews a bank's balance sheet, mistakenly judges the bank is insolvent, and publicly predicts that it will fail. When depositors learn this, it affects their beliefs about the bank's future, this feeds back on their behavior in the present causing them to withdraw their deposits, and then the bank indeed fails, making the analyst's false belief come true.

To see how two sets of expectations are involved, notice that the bank analyst's judgment about the bank's solvency is ultimately a judgment that the bank's stock price is too high, whereas depositors' expectations are rather about getting their money back and thus really about the bank's underlying fundamentals. At most, depositors are worried about the bank's stock price as an indicator of the safety of their money. However, the bank run directly affects the bank's fundamentals, and this then feeds over to and changes investors' expectations about the bank's stock price. The dynamic between the two sets of expectations consequently works through how a change in expectations about the bank's fundamentals influences expectations about its stock price, not the other way around. It might seem that the turning point is the analyst's judgment about the bank's stock price, and that investors are in the lead. But this initially has no effect on investors. Indeed, the advantage of Merton's example is that it specifically rules this out in that the bank analyst makes what is generally regarded as a false judgment about the bank according to the standards of the financial community, which had instead judged the bank to be solvent. It is only after the bank run and change in depositor expectations that investors judge the bank stock price to be too high. Thus, at the turning point it is the reversal in beliefs about fundamentals, not about the bank's stock price, that leads to a reversal in beliefs about the bank's value and precipitates the downswing.

⁴ Self-fulfilling prophecies are earlier referred to in social science as the Thomas theorem: 'If men define situations as real, they are real in their consequences' (Thomas and Thomas, 1928, pp. 571-2).

This self-fulfilling prophecy turning point explanation, then, also describes the 2007-2008 U.S. banking crisis. The crisis also essentially began with a bank run after which the stock market capitalization of the major U.S. banks declined dramatically, Bear Stearns and Lehman Brothers failed, and most large banks had to be bailed out because inter-bank lending dried up (Brunnermeier, 2009). Up until the crisis, most speculators and investors regarded the banks as solvent, and dissenters were thought to be wrong by most analysts in the financial community, as well described by Michael Lewis (Lewis, 2010). Yet the dissenters eventually influenced some investors who began to be more cautious. Inter-bank lending then began to decline, banks found they were no longer able to finance themselves on overnight markets, and were then illiquid and in some cases insolvent, fulfilling the dissenters' 'false' prophecy. As in Merton's example, then, it was a reversal in beliefs about the banks' fundamentals, not beliefs about bank stock market values, which led to a reversal in beliefs about the value of the banks and brought about the downswing. In both Merton's example and the case of the banking crisis, then, expectations that were 'out of balance' in the upswing when speculation drove fundamentals were rebalanced at the turning point when expectations regarding fundamentals took the lead.

ii. Self-defeating prophecies

Consider now the symmetrical case of a self-defeating prophecy, where instead of something false becoming true, something regarded as true becomes false.⁵ A self-defeating prophecy involves a correct judgment and prediction that sets in motion behavior that overturns it. The paradigmatic example often cited is the Y2K prediction, where the prediction that the year 2000 calendar turnover would cause all computers to crash led governments, businesses, and individuals to act in ways that assured it would not occur and thus made the prediction false. Of course, economic agents always have a combination of true and false beliefs, and thus both can be 'tested' and overturned in their interaction with one another. In a general crisis, then, in which there are multiple, interconnected boom-bust cycle turning points and a wholesale reversal in beliefs, we need to understand how true beliefs can be reversed along with false ones. Self-

⁵ Self-defeating prophecies are also referred to as the Oedipus effect after Oedipus' legendary efforts to avoid the prophecy about his future actually helped bring them about.

defeating prophecies are also especially important to understand, because they go against people's ordinary philosophical intuitions are that what is true cannot not be true.⁶

The Y2K case is an example of a self-defeating prophecy in which policies were undertaken to prevent a crisis. However, self-defeating prophecies can also produce crises and turning points in boom-bust cycles, as can be argued in the connection with the U.S. housing crisis. Like Merton's bank run case, we begin with judgments that occur on the financial side of the story. Thus, most households and analysts of the housing sector were aware households were becoming more leveraged in the late 1990s and early 2000s but believed that the asset side of their balance sheets was improving with higher home prices, thus offsetting their expansion on the liability side due to higher mortgage finance requirements. This judgment was made by the firms that were lending to households, in the business media that continually claimed housing prices would never fall because they had never fallen, by banks which sold the associated derivative products to investors and one another on the model that derivative assets were superior to conventional mortgage assets, by credit rating agencies who reviewed the banks, and by the U.S. Federal Reserve which denied that bubbles were an issue in financial markets. In hindsight, these judgments were mistaken, but the point is that they were thought to be true when they were made.

The effect of these judgments, however, was to increase complacency towards risk throughout the financial system in classic boom style, which meant that all the processes involved continued until households had indeed taken on more debt than their incomes could justify. This then created doubts over whether those households were in a position to finance their mortgages; this

⁶ Self-defeating prophecies are sometimes confused with self-refuting hypotheses. Self-defeating prophecies are predictions with causal effects. Self-refuting hypotheses are claims or arguments said to logically contradict themselves. For example, Burton Malkiel used the idea to characterize stock price movements as a random walk in defense of the efficient markets hypothesis. Technical forecasting methods that aim to identify trends in stock market prices, he argued, must be self-defeating: "The problem is that once such a regularity is known to market participants, people will act in such a way that prevents it from happening in the future" (Malkiel, 2003, p. 168). Sanford Grossman and Joseph Stiglitz's argument that the idea of informationally efficient markets is impossible can also be regarded as an example of a self-refuting hypothesis: if markets did convey all possible information, agents would have no reason to acquire the information necessary to make them efficient (Grossman and Stiglitz, 1980). A classic type of self-refuting hypothesis is a self-referential claim that contradicts itself, such as the liar paradox: Epimenides, a Cretan, asserted that all Cretans are liars.

slowed mortgage finance initially to the subprime sector, and this slowed the housing price rise; slowing price increases then meant that other households would also be less likely to be able to finance and re-finance themselves, which further slowed the lending fueling the price rise; it then indeed became false that household balance sheets were secure, thus transforming what many formerly believed to have been a true judgment into one that had become false. Consequently, as in the self-fulfilling prophecy case, here in the case of a self-defeating prophecy it was a reversal in beliefs about fundamentals – in this instance, household fundamentals – that led to a reversal in beliefs about household balance sheets that brought about the downswing in home prices.

iii. From turning points to booms and upswings

The argument above, then, is that the phase transition at the turning point in a cycle occurs when expectations about fundamentals shift relative to expectations about stock prices, rebalancing them toward one another after they had become increasingly ‘out of balance’ in the upswing. This occurs in those exceptional circumstances at the peak of a cycle when a prediction, true or false, is made about the financial status or stock price of some agent or agents, and precipitates a wave of belief reversals about economic fundamentals, which precipitates a further wave of belief reversals about the financial status or stock price, which together then have the effect of reversing the truth of the prediction.

However, of course most predictions are unlike self-fulfilling and self-defeating prophecies, and do not generate widespread belief reversals that reverse their truth. That is, most predictions are benign and belief-preserving rather than belief-reversing, and accordingly people are more accustomed to understanding prediction in this less dramatic, more ordinary way. Soros’ reflexivity analysis brings this out by treating the ‘far-from-equilibrium’ boom-bust cycle under positive feedback as a departure from the ‘near-equilibrium’ conditions driven by negative feedback that prevail most of the time. When we turn to the boom and upswing, then, the misconception that he believes agents have in a boom also derives from its being a departure from their ordinary view of the world when prediction is relatively reliable. In the following section, to explain this misconception I trace it to how the way people ordinarily reason about cause-and-effect in connection with feedback processes in general causes them to misunderstand the nature of positive feedback processes in the upswing of a boom in particular.

4. Second argument: How expectations get ‘out of balance’

Reflexivity theory assumes that if people are able to adjust their behavior in light of their expectations of the future, they must understand something about the nature of cause-and-effect processes in the world in which they live. Otherwise it would make no sense for them to change what they do in the present in an effort to change how the future plays out. In addition, it is also fair to assume that what people understand about cause-and-effect processes depends on the evidence they have about them from the past. Accordingly, I suppose that most economic agents operate with a very basic philosophy of science reasoning whereby they more or less consciously hypothesize cause-and-effect models about economic processes in light of their observation of repeated feedback patterns. The problem with this, however, is that this conception only fits negative and some positive feedback patterns but does not fit those positive feedback patterns we observe when the economy is undergoing significant change and the past is an unreliable guide to the present. This then leads agents to misconceive positive feedback patterns in booms when the economic relationships are likely undergoing significant change and will not be stable, and it is this that causes expectations about stock prices to get increasingly ‘out of balance’ with expectations about fundamentals. To develop this argument, first, I discuss how economic agents are likely to reason about feedback patterns when they model cause-and-effect relationships, second I discuss the often neglected asymmetry between negative and positive feedback patterns, and third I discuss how these things together generates a misconception about booms that causes expectations about stock prices to get increasingly ‘out of balance’ with expectations about fundamentals.

i. Feedback patterns as repeated reflexivity judgments

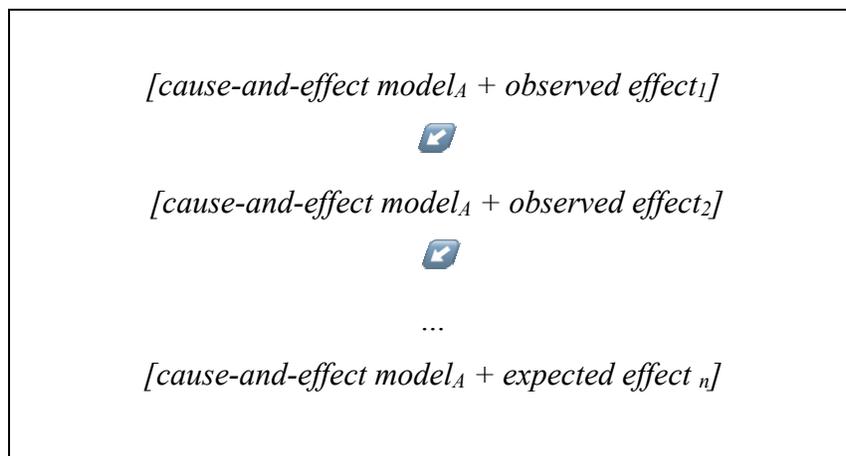
Reflexivity was defined in basic terms in section two without saying how a sequence of reflexivity judgments might follow some repeated feedback pattern. The basic idea there was that people’s views of the future feed back onto and *somehow* influence their behavior in the present, which then *somehow* affects how the future plays out, which then *somehow* affects their subsequent views of the future, and so on, as in Figure 1. Feedback patterns, however, are

repeated sequences of reflexivity judgments of some regularity that economic agents observe, identify, and rely upon to hypothesize the existence of cause-and-effect relationships they can act upon.

Consider the relatively predictable positive feedback case of firms that have decreasing costs. Here the view that firms have of the future or what they expect is that increasing production will cause unit costs to fall. When this expectation feeds back onto the present and firms in fact increase current production, this affects how the future plays out by causing prices to fall, and if this pattern then occurs repeatedly, this affects these firms' view of the future as one in which they can rely on an empirically well identified cause-and-effect relationship between increased production and lower costs.

Thus, a feedback pattern involves a settled sequence of reflexivity judgments that provides a basis for a hypothesized cause-and-effect model for economic agents. Of course, models must be hypothesized because they represent unobservable cause-and-effect relationships underlying our observations about feedback patterns. I represent this kind of reflexive reasoning in Figure 2 as it pertains to a single cause-and-effect model A . The downward arrows indicate that economic agents believe cause-and-effect model $_A$ is repeatedly confirmed, because the sequence of observed effects $1, 2, \dots, n$ for the model are consistent with what the model predicts. In this case, agents thus make successive belief-preserving judgments, and are confident that the model predicts reasonably well.

Figure 2: Ordinary cause-and-effect reasoning about feedback patterns



Should, however, the evidence not support a model, what happens? One response agents are likely to have is to suppose that cause-and-effect model *A* should be replaced by some other cause-and-effect model *B* and that a new Figure 2 applies. That is, if the evidence changes, the model must be changed. This is what happens in Merton's example of the bank analyst making an influential judgment about a bank. In that case, agents had held a cause-and-effect model in which the bank's continuing to take in deposits for the foreseeable future was the cause of the bank remaining solvent in the future. The bank analyst's judgment, however, led depositors to expect a bank run, which meant the cause-and-effect model that predicted the bank would remain solvent no longer applied. This leads to their adoption of another model based on another sequence of reflexivity judgments involving an alternative feedback pattern. Depositors now expect insolvency in the future, this feeds back onto their behavior in the present by increasing current withdrawals, which affects how the future plays out in terms of the worsening condition of the bank, which reinforces expectations about the future, and so on in a new feedback pattern that identifies an alternative cause-and-effect model that predicts the bank's collapse.

Yet while this reasoning about cause-and-effect models makes sense for the firm with decreasing costs and the Merton example where the fit between past evidence and the corresponding models is often good, it ignores the possibility that this sort of reasoning about cause-and-effect models is inappropriate in circumstances in which economic relationships are in a process of change, unstable, and largely in flux. To see this when this can be the case, note how negative and positive feedback patterns differ in the ways in which they relate causes and effects to one another.

ii. Negative versus positive feedback patterns

Negative feedback, or diminishing returns, occurs where some sort of homeostatic or self-regulating process is operating in which cause-and-effect relationships are sustained in largely unchanged form over time. As a repeated pattern, negative feedback occurs when successive increases or decreases in an activity are always decreasing due to a feedback loop from effects to

causes that dampens the force of the causes. In the classic example of diminishing returns in production, the feedback loop works through the decreasing productivity or reduced causal effectiveness of higher input levels on output. The process is self-regulating in the sense that its primary effects – increased output in this case – are always offset by its secondary effects – reduced incentive to further increase output.

Negative feedback is thus fully consistent with the ordinary cause-and-effect reasoning of Figure 2. Because the repeated effects of any activity always increase at a decreasing rate, economic agents can reasonably infer that nothing is happening that counts as evidence that they have the wrong cause-and-effect model. This conclusion can be established through a process of indirect reasoning. If the activity were rather increasing at an increasing rate and its causal effects continually growing, it would make sense to suppose something is additionally causing those effects beyond the causes originally assumed. The absence of increasing effects, however, removes this concern, and this means there is no reason to question the way in which cause and effect are connected in the model.

To illustrate, imagine that as a bank expands its balance sheet, withdrawals increase but at a decreasing rate. The larger size of the bank implies a higher level of withdrawals, but if this occurs at a decreasing rate, the cause-and-effect model for the bank implies solvency. If, however, the expansion of the bank's balance sheet leads withdrawals to increase at an increasing rate, many would worry that some additional set of causal forces is operating that affects withdrawals, and thus that their cause-and-effect model of the bank as solvent might be wrong. But this does not occur. What makes negative feedback patterns consistent with Figure 2, then, is that agents do not any have a reason to think the cause-and-effect model they have identified from repeated feedback patterns is wrong.

Consider now positive feedback, or increasing returns, which involve successive increases or decreases in an activity that are increasing. When distinguished from one another in terms of secondary effects, negative and positive feedback processes appear to be simply the opposite of one another. In fact, they are asymmetrical in how they relate causes and effects to one another. For negative feedback, because the repeated effects of any activity always increase at a decreasing rate, such that the repeated effects of that activity always get smaller and smaller, economic agents can reasonably infer that nothing is happening on the cause side of the

relationship to overturn their cause-and-effect model. However, when the effects of an activity are always increasing, there is a possibility that something additional is occurring on the cause side of the relationship, thus implying that the existing cause-and-effect model may no longer apply. The original model *might* still apply, and the increasing effects *might* not mean there is anything new happening on the cause side of the relationship. Nonetheless the possibility something new is happening cannot be ruled out. There is thus an important difference between the two forms of feedback in that increasing effects opens the door to the possibility that new causal forces are operating, whereas negative feedback and the diminishing effects it involves firmly closes the door to it.

At the same time, it does not follow from the idea that ‘new causal forces’ might be operating that we simply need to replace cause-and-effect model *A* with some other cause-and-effect model *B*. In the reflexive social world, we occupy, people’s actions change the world. This invests change with an ontological status and means that the underlying relationships we investigate cannot always be in a settled state. Economic relationships can of course be stable and relatively unchanging, but they can also simply be changing and in flux, in which case they are not subject to explanation employing basic philosophy of science reasoning. Thus when we address examples of positive feedback processes we do well to distinguish these forms, or better, think in terms of a spectrum of cases, where at one the end of the spectrum there are cases which do not involve significant change in economic relationships, such as (arguably) firms with decreasing costs increasing their shares of a market but with production and demand schedules that are still reasonably predictable, and where at the other end of the spectrum there are cases involving wholesale change in economic relationships, such as such as when new technologies are introduced that have transformative effects on many markets and economic relationships (Arthur, 1994).

Of course, these two ends of the spectrum differ in degree, and individual cases can surely be debated as to where they fall on the spectrum since ‘change in economic relationships’ can include so many different types of considerations. However, that there can be such a debate – indeed that there must be such a debate - is central to what is important about positive feedback compared to negative feedback, because it means there exists an inherent ambiguity about what any instance of positive feedback involves. This kind of ambiguity is different from that

associated with the standard problem of hypothesizing cause-and-effect models from the evidence we have about them. In the case of positive feedback processes falling at the transformational end of the spectrum, we may have a body of evidence about them which may even exhibit certain regularities, but we should infer that no model conceived of in terms of a stable set of underlying relationships can represent that evidence. That is, our basic philosophy of science reasoning about cause-and-effect gives an inadequate explanation of the world when we encounter positive feedback processes that fall at the transformational end of the spectrum.

Nonetheless, the standard view has an undeniable appeal as a mode of reasoning. Given this, we should not be surprised if economic agents rely on it and are systematically mistaken in booms when positive feedback appears to govern economic fundamentals. Then, when they find that the evidence for their views is no longer supported, we should expect them to quickly lose confidence in those views. Let us use this analysis, then, to explain agents' misconceptions in the upswing of a boom.

iii. Misconceptions in the upswing of a boom

In the upswing of a boom, positive feedback processes for both stock prices and economic fundamentals generate a dynamic of interconnected sets of expectations regarding stock prices and fundamentals. Speculative expectations raise stock prices at an increasing rate, and this causes economic fundamentals in the form of earnings per share to improve at an increasing rate. This then influences expectations about fundamentals, so what happens is that the change in expectations about stock prices has changed expectations about fundamentals, at least until the turning point in the cycle when the influence runs in the opposite direction. At that point – the 'Minsky moment' – prior expectations about fundamentals re-assert themselves and pull down expectations about stock prices. When the downswing has run its course, 'near-equilibrium' conditions and negative feedback processes again determine expectations for stock prices and fundamentals, and as fundamentals only increase at a decreasing rate determined by long run factors, both sets of expectations adjust accordingly. Essentially, then, what the 'far-from-equilibrium' experience of the boom does is create a short-term misconception about fundamentals that temporarily allows speculative expectations to run their course until a realism about fundamentals ends the boom.

Why, then, do economic agents become mistaken about fundamentals during the boom so that their expectations adjust to traders' speculative expectations? Expectations about fundamentals are grounded in economic performance whereas speculative expectations depend on risky bets and trader psychologies. The former are dominant in the normal course of events before and after booms, but oddly during the boom expectations about fundamentals depart from what these 'near-equilibrium' conditions support and chase speculative expectations. On the argument above, this happens because economic agents believe that the positive feedback processes they are observing regarding fundamentals in the upswing represent new, stable cause-and-effect relationships improving earnings per share, when it is rather the case that the underlying cause-and-effect relationships involved are unstable and changing, and do not justify this belief. Economic agents reason in a traditional philosophy of science manner, and therefore suppose that there *must* be some underlying cause-and-effect model to explain the positive feedback patterns they observe. Accordingly, when doubts emerge about their models, they have no way of explaining the situation, and default to their better confirmed *status quo ante* views of fundamentals.

We can illustrate this by what happened in the 2007-2008 U.S. housing crisis. The housing market in the 1990s and early 2000s exhibited positive feedback processes in both home prices and expansion in household balance sheets beginning in the subprime market and eventually extending to all housing. Whatever caused prices to start rising in the first place, once they did speculative expectations developed, and this led to increasing lender finance. This improved household fundamentals, and thus favorably changed expectations about households' capacity to finance more debt, so that speculative expectations led to adjustment in expectations about household fundamentals. The theory that economic agents had, then, was that this was a 'new' world in which expanded household balance sheets were secure, because home prices rising at an increasing rate justified household leverage rising at an increasing rate. What was empirically observed, then, provided grounds to hypothesize a new underlying cause-and-effect model of the housing market that replaced the old banking model of traditional household finance that had prevailed previously under conditions of negative feedback.

Yet as later became clear, what was actually occurring in household finance did not reflect a new stable set of relationships between housing debt and household income. Rather the situation was

fully in flux as households used increased access to finance to offset incomes that had been weakening since the 1980s. If household income had not been declining, the observed positive feedback story might have reflected a stable new set of relationships, if it could be argued that the fundamentals for the long-run value of U.S. housing had improved. Then the positive feedback pattern could have explained a short-term adjustment until new long-run values were established. However, it ultimately became clear that the situation was not stable and household debt was rising unsustainably, and confidence regarding there being a new underlying model for housing then disappeared. The complacency agents exhibited about there being a new underlying model for housing worked as a self-defeating prophecy that led to further increases in lending until expectations regarding household fundamentals declined, which then pulled down speculative expectations over house prices.

Thus, the misconception economic agents were under was that they thought the world was working predictably despite the evidence that it was changing. They were essentially blind to that evidence because they implicitly employed a philosophy of science reasoning regarding causality that emphasizes predictability. Indeed they employed a natural science type of causal reasoning rather than a social science type of causal reasoning emphasizing reflexivity. This leaves unexplained, however, why economic agents are vulnerable to this misconception. Why is their philosophy of science reasoning overly simplistic in too quickly underpinning observed empirical regularities with hypothetical cause-and-effect models?

The answer, I suggest, lies in how well that reasoning fits their experience much of the time when negative feedback patterns prevail. Negative feedback is always consistent with there being some cause-and-effect model fitting Figure 2, and so when economic agents conclude they have the wrong underlying cause-and-effect model, they are inclined to suppose that what they need to do is find the right model – not to think that no such model might exist. This inclination is reinforced by the fact that boom-bust cycles constitute a departure from the normal, ‘near-equilibrium’ course of events that the basic model often explains quite well. Quite simply, then, economic agents are biased toward seeing the world in terms of negative feedback processes, and this blinds them to the need to be able to explain positive feedback processes associated with transformational change in the economy.

5. Boom-bust cycles and the nature of reflexive economic agents

This paper developed a reflexivity understanding of the interdependent expectations of interacting adaptive economic agents in financial markets in connection with boom-bust cycles. Reflexivity operates through a feedback loop from how people's expectations of their future interaction feed back on and influence their current actions and strategies to how their current actions and strategies feed forward on and influence their future interaction. When we think of this in terms of two interacting sets of expectations, we can explain boom-bust cycles dynamically according to how these feedback processes cause these different sets of expectations to change in relation to one another across the cycle as a single evolving structure of interaction.

The argument in the paper first addressed turning points, because they reveal the structure of interaction in a set of interdependent expectations in a boom-bust cycle, but the second the argument on the misconception agents have about the upswing in the cycle goes to a deeper issue, namely, why agents are blind to the causes of the cycle. Though the answer may seem to be a behavioral economics type explanation, it is less about how people perceive the world, and more about how stability in economic relationships gives precedence to a basic philosophy of science view.

Central to the whole analysis, in any case, is paper's conception of reflexive economic agents seen as adaptive and interactive. As adaptive, their behavior needs to be explained in terms of how they are embedded in time; as interactive, their behavior needs to be explained in terms of how they are socially embedded. These two characteristics contrast with the mainstream atomistic agent view that rejects both characteristics. Not surprisingly, mainstream economics itself is neither complex nor evolutionary. Not surprisingly, the mainstream framework fails to explain recurring crises and financial fragility that are so disruptive and costly to people's lives. Thus, one motivation for this paper's argument is its assumption that adequate explanations of capitalist market economies depend on first having adequate conceptions of economic agents who inhabit such economies. This paper makes one step in that direction in connection with its analysis of boom-bust cycles.

Acknowledgements: The author is grateful for comments on a previous version of this paper to David Colander, Wolfram Elsner, and Wade Hands, for discussions regarding reflexivity to Nicolas Brisset, Barkley Rosser, and George Soros, and to Brian Arthur, Eric Beinhocker, and Alan Kirman for assistance in thinking about complexity.

Disclosure statement: No potential conflicts of interest were reported by the author.

Open access declaration: This paper does not require open access publication.

References

Arthur, Brian (2013) “Complexity Economics: A Different Framework for Economic Thought,” SFI Working Paper 2013-04-12, Santa Fe Institute.

Arthur, W. Brian (1994) *Increasing Returns and Path Dependence in the Economy*, Ann Arbor, MI: University of Michigan Press.

Brunnermeier, Markus (2009) “Deciphering the Liquidity and Credit Crunch,” *Journal of Economic Perspectives* 23 (1): 77-100.

Davidson, Paul (2009) *The Keynes Solution: The Path to Global Economic Prosperity*, London: Palgrave/Macmillan.

Davis, John (2018) “Extending Behavioral Economics’ Methodological Critique of Rational Choice Theory to Include Counterfactual Reasoning,” *Journal of Behavioral Economics and Policy* 2 (2).

Davis, John (2017) “The Continuing Relevance of Keynes’s Philosophical Thinking: Reflexivity, Complexity, and Uncertainty,” *The Annals of the Fondazione Luigi Einaudi. An Interdisciplinary Journal of Economics, History and Political Science*, LI 1 (2017) 55-75.

Grossman, Sanford and Joseph Stiglitz (1980) “On the Impossibility of Informationally Efficient Markets,” *American Economic Review* 70 (3): 393-408.

Keynes, J. M. (1936) *The General Theory of Employment, Interest and Money*, London: Macmillan.

Kirman, Alan (2010) “The Economic Crisis is a Crisis for Economic Theory,” *CESifo Economic Studies* 56 (4): 498-535.

Lewis, Michael (2010) *The Big Short: Inside the Doomsday Machine*, New York: W. W. Norton & Company.

Malkiel, Burton (1973) *A Random Walk Down Wall Street*, New York: W. W. Norton & Company.

Marchionatti, Roberto (2010) “J. M. Keynes, Thinker of Economic Complexity,” *History of Economic Ideas* 18 (2): 115-146.

Marcuzzo, Maria Cristina (2020) “Pets and favorites: Keynes practice as investor in the stock exchange,” in W. Dolfsma, D. W. Hands, and R. McMaster, eds., *History, Methodology and Identity for a 21st Century Social Economics*, Abington: Routledge.

Merton, Robert (1949) *Social Theory and Social Structure*, New York: Free Press.

Minsky, Hyman (1982) *Can 'it' Happen Again? Essays on Instability and Finance*, Armonk, NY: M. E. Sharpe.

Samuelson, Paul (1993) "Reflections on Investing for Foundations and Colleges," Speech at Rutgers, Nov. 17, 1993, published in: *The Collected Scientific Papers of Paul A. Samuelson*, Volume 7, Janice Murray, ed., Cambridge, MA: MIT Press, 2011: 151-157.

Soros, George (2013) "Fallibility, reflexivity, and the human uncertainty principle," *Journal of Economic Methodology* 20 (4): 309-329.

Soros, George (1987) *The Alchemy of Finance*, Hoboken, NJ: Wiley & Sons.

Thomas, William and Dorothy Thomas (1928) *The child in America: Behavior problems and programs*, New York: Knopf.