DISS. ETH NO. 27828

Accelerating Innovation:
Markets, Speculation, and Emerging Technologies

A thesis submitted to attain the degree of DOCTOR OF SCIENCES of ETH ZURICH (Dr. sc. ETH Zurich)

presented by

TOBIAS A. HUBER

MA, Philosophy, UWE Bristol
born on 29.07.1985

citizen of Switzerland, place of origin: Uster, Zurich

accepted on the recommendation of
Prof. Didier Sornette
Peter Cauwels

2021

Abstract

This cumulative dissertation examines the nature and dynamics of technological innovation. By drawing from the economics of innovation, financial economics, and the philosophy and sociology of science and technology, the dissertation addresses in three case studies the question of what drives progress in technology and science. The aim of this project is to contribute to a general theory of innovation that is centered around the role of bubbles in the development and diffusion of emerging technologies.

The first research paper aims at developing a deeper understanding of scientific progress. In a case study that focuses on the evolution of financial economics, it dissects the historical cross-fertilization between physics and finance. Based on an analysis of how theories and models get constructed in the emerging field of econophysics, we show how cutting-edge insights from physics have substantially improved the scientific understanding of financial markets. We conclude the paper by arguing that a new science of economic and financial systems needs to integrate findings from other scientific fields so that a truly multi-disciplinary complex systems science of financial markets can be built.

The second research paper advances the Social Bubble Hypothesis, a powerful conceptual framework that has been developed in our research group to illuminate the mechanisms underlying scientific and technological innovation. The Social Bubble Hypothesis holds that bubbles constitute an essential component in the process of technological innovation and adoption. In this paper, we apply the Social Bubble framework to Bitcoin. We examine the role of speculative bubbles in the process of Bitcoin's technological adoption by analyzing its social dynamics. The paper traces Bitcoin's genesis and dissects the nature of its techno-economic innovation. In particular, it presents an analysis of the techno-economic feedback loops that drive Bitcoin's price and network effects and reveals how a hierarchy of repeating and exponentially increasing series of bubbles and hype cycles has bootstrapped Bitcoin into existence.

The third research paper further extends the Social Bubble Hypothesis by applying it to the clean-tech bubble that formed between 2004 and 2008 in clean technologies, such as solar, biofuels, and batteries. The paper synthesizes the development of the clean-tech bubble, its history, and the role of venture capital and government funding in catalyzing it. In particular, we dissect the underlying narratives that fueled the bubble. The paper concludes by presenting evidence that the clean-tech bubble constitutes an example of an innovation-accelerating bubble.

Kurzfassung

Diese kumulative Dissertation untersucht das Wesen technologischer Innovation. Sich auf die Ökonomie der Innovation und die Philosophie und Soziologie der Wissenschaft und Technologie beziehend, geht sie in drei Fallstudien der Frage nach, was den Fortschritt in Technologie und Wissenschaft antreibt. Ziel dieses Projekts ist es, einen Beitrag zu einer generellen Theorie der Innovation zu leisten, in deren Mittelpunkt die Rolle von spekulativen Blasen in der Entwicklung und Adoption neuer Technologien steht.

Das erste Forschungspapier zielt darauf ab, ein tieferes Verständnis des wissenschaftlichen Fortschritts zu entwickeln. In einer Fallstudie, die sich auf die Entwicklung der Finanz-Wissenschaft konzentriert, wird die historische wechselseitige Befruchtung zwischen Physik und Ökonomie untersucht. Auf der Grundlage einer Analyse der Art und Weise, wie Theorien und Modelle in der Ökonophysik konstruiert werden, zeigen wir, wie bahnbrechende Erkenntnisse der Physik das wissenschaftliche Verständnis der Finanzmärkte wesentlich verbessert haben. Abschließend argumentieren wir, dass eine neue Wissenschaft der Finanzsysteme Erkenntnisse aus anderen Wissenschaftsbereichen integrieren muss, damit eine wahrhaft multidisziplinäre Complex-Systems Wissenschaft der Finanzmärkte entwickelt werden kann.

Das zweite Forschungspapier stellt die Social-Bubble-Hypothese vor, die in unserer Forschungsgruppe entwickelt wurde, um die Mechanismen zu beleuchten, die wissenschaftlicher und technologischer Innovation zugrunde liegen. Die Social-Bubble-Hypothese postuliert, dass Blasen eine wesentliche Komponente im Prozess der technologischen Innovation darstellen. In diesem Papier wenden wir die Social-Bubble-Hypothese auf Bitcoin an. Wir untersuchen die Rolle von spekulativen Blasen im Prozess der technologischen Adoption von Bitcoin, indem wir die soziale Dynamik von Bitcoin analysieren. Das Papier zeichnet die Entstehungsgeschichte von Bitcoin nach. Insbesondere werden die techno-ökonomischen Feedback-Loops analysiert, die den Preis und die Netzwerkeffekte von Bitcoin antreiben. Wir zeigen zudem auf, wie eine Hierarchie sich wiederholender und exponentiell-zunehmender Blasen und Hype-Cycles Bitcoin antreibt.

Das dritte Forschungspapier erweitert die Social-Bubble-Hypothese, indem wir sie auf die Clean-Tech-Blase anwenden, die sich zwischen 2004 und 2008 in Clean Technologien wie Solar, Biokraftstoffen oder Batterien gebildet hat. Das Papier fasst die Entwicklung der Clean-Tech-Blase, ihre Geschichte und die Rolle, die Risikokapital und staatliche Finanzierung in ihrer Entstehung gespielt haben, zusammen. Insbesondere untersuchen wir die zugrundeliegenden Narrative, die die Blase angeheizt haben. Abschließend wird der Nachweis erbracht, dass die Clean-Tech-Blase ein Beispiel für eine innovations-beschleunigende Blase ist.

Acknowledgements

This PhD thesis has profited from the support and feedback of numerous people and institutions. First of all, I wish to express my gratitude to Professor Didier Sornette for being a source of invaluable inspiration. His commitment, relentlessness, and open-mindedness made this thesis possible. I would also thank Peter Cauwels for serving as a co-examiner and Judith Holzheimer for her assistance. I also thank my colleagues for productive debates—especially, Spencer Wheatley, Vladimir Filimonov, Richard Senner, Ke Wu, Zhuli He, and the members of the Complexity Economics Working Group at the Institute for New Economic Thinking. I am also very thankful to Byrne Hobart for providing valuable feedback and input.

Most importantly, I would like to thank my wife. Without her, I would not have been able to complete this project. I am grateful for my parents' support. I dedicate this thesis to my family.

Contents

1. Introduction

Humanity depends on technological and scientific progress. Over the last 250 years, all civilizational advances and socio-economics transformations have been fundamentally shaped by techno-scientific innovation. However, even though there is an expanding literature, the nature of technological innovation and its driving factors are still intensely debated. Given that the future of our techno-economic and socio-political systems is premised on ever-accelerating progress, the question concerning the nature of scientific and technological innovation can thus be considered to be one of the most important questions of our age. It is therefore critical to develop a deeper understanding of the essence of techno-scientific innovation.[1] By drawing from the economics of innovation and the philosophy and sociology of science and technology, this thesis seeks to illuminate the nature of techno-scientific progress in a series of case studies that examine the structure and dynamics of innovation. The aim of this project is to contribute to a general theory of innovation that is centered around the role of bubbles in the development and diffusion of emerging technologies.

Over the past decades, three dominant theories of progress have emerged: one view, more techno-utopian in nature, holds that uncontrollable and irreversible technological growth will super-exponentially increase until we reach the "technological singularity"—"an opaque wall across the future" (Vinge 1993) beyond which we will witness the creation of a "super-intelligence" in which human and machine merge (see Kurzweil 2005; Bostrom 2014). The Singularitarian view contrasts with the catastrophic vision that economic growth and technological progress will inevitably result in civilizational collapse or environmental crisis (see Meadows and Randers 2012). A third view, which has become more popular over the past decade, holds that—contrary to our apparent experience of ever-intensifying innovation—technological innovation is stagnating: instead of asymptotically approaching a technological singularity, we have, according to this view, entered a paradigm of decaying economic growth and increasing social disorder.[2] Most prominently developed

---

[1] While this thesis aims to identify recurring patterns in technological innovation, it does not subscribe to an essentialist conception of technology, which conceptualizes technology as following an inhuman internal logic that transcends cultural and social processes. For an example, see Heidegger's ontological conception of technology that cannot be explicated in terms of instantiations of specific technologies (Heidegger 1990).

[2] "Innovation" is a multifaceted concept. Cauwels and Sornette (2020) provide a useful taxonomy that distinguishes "innovation" from "invention" and "discovery." In their classification, "discovery" refers to an "observation or recognition of something already existing," innovation denotes the "creation of something new," and "innovation" signifies the recombination, optimization, or incremental improvement of an existing technology. Now, in order to test the hypothesis that a slowdown in the rate of technological and scientific discoveries is occurring, Cauwels and Sornette propose that it is important to distinguish whether a new idea stems from a discovery, an invention, or an innovation. Based on the "Flow of Ideas" and the "Research Productivity" metrics—which they identify as the two fundamental drivers of economic growth—they construct a series of indices that quantify the historical evolution of fundamental scientific knowledge. Whereas the "Flow of Ideas" indices give the number of discoverers and inventors from different scientific disciplines and geographical areas, who are active in their professional career, for each year, the "Research Productivity" indices

by economists Tyler Cowen and Robert Gordon and investor Peter Thiel, the Stagnation Hypothesis locates the onset of techno-economic decline in the early 1970s (see Thiel 2011; Cowen 2011; Gordon 2016). Proponents of the Stagnation Hypothesis mobilize as evidence stagnating wage growth, increased inequality and income concentration, an explosive increase in debt and leverage, massive increases in the consumer price index, or the halving of Total Factor Productivity, which measures growth in aggregate output not explainable with an increase in labor and capital inputs (see Sornette and Cauwels 2014).

While it is beyond the scope of this thesis to evaluate the empirical adequacy of these theories of progress, it rejects the fundamental assumption that is embedded in the views that innovation will either inevitably lead to a technological singularity or economic collapse: namely, that progress is an automatic or inevitable process. On the contrary, against these more deterministic or quasi-teleological conceptions of progress, this thesis assumes that progress is contingent on human agency and a specific set of contextual and, often irreducibly contingent, path-dependent factors. Consequently, the methodology of this thesis seeks to reverse-engineer how transformative progress arises based on a series of case studies. While it focuses on concrete cases, the aim of this thesis is nevertheless to extract generalizable patterns, features, and factors that make technological innovation or breakthroughs in technology, engineering, and science possible. By identifying the constitutive elements of innovation, this thesis makes it possible to infer the outlines of a general theory of technological progress.

1.2 Models of Progress

Over the past centuries, a series of philosophical models of progress have been proposed.[3] One of the most dominant models, which informs this thesis, is the model of the structure of scientific revolutions that has been developed by philosopher and sociologist of science Thomas Kuhn and its extension and application to the dynamics of technological innovation by economic historian Carlota Perez.

At their core, both conceptual models posit that innovation—in science and technology—progresses through revolutions or "paradigm-shifts." In the Kuhnian framework, the steady and continuous process of scientific development, which Kuhn calls "normal science" and that he characterizes as "puzzle-solving," is sporadically disrupted by scientific revolutions (see Kuhn 2012). Revolutions, in

provide the per capita (per million) result of the first series. Cauwels and Sornette observe a "consistent decline [...] across the board since the early 1970s, both for the 'Flow of Ideas' and the 'Research Productivity'" metrics. Furthermore, an interesting finding is that the "digital revolution" does not appear to reverse the decline in technological progress, which, as the authors conclude, might suggest that "the digital revolution may be mainly driven by innovation and exploitation of existing knowledge, and not by discovery and invention, or new explorations" (see Cauwels and Sornette 2020).

[3] Famous examples include Hegel and Marx' teleological models of history or Spengler's cyclical model of civilizational progress (see Hegel 2007; Marx 1993; Spengler 1991, Strauss 1981). For a critique, see Popper (2013) and Hayek (1952).

this model, are catalyzed by the emergence of "anomalies," which are incompatible with the existing paradigm that dominates normal science. A scientific revolution follows a crisis when a novel paradigm, which is "incommensurable" with the practice of normal science, supersedes the previous paradigm.[4] Expanding Kuhn's model of progress, socio-economist Carlota Perez develops the concept of "technological revolutions" to explain discontinuities and regularities in the historical process of innovation (see Perez 2003). Parallelizing the Kuhnian view on the nature of scientific progress, Perez identifies a sequence of technological revolutions and "techno-economic paradigms" that have disrupted our industries and societies over the past three centuries.

In her work on the economics of innovation and technological change, a technological revolution—which locally disrupts a specific market or industry in terms of new inputs, methods, and technologies—becomes a techno-economic paradigm when it starts to globally transform organizational structures, business models, and strategies in markets and sectors beyond which the technological breakthrough had initially erupted. Techno-economic paradigms, in other words, represent a collectively shared best practice model of the most successful and profitable uses of new innovations. By enabling the wide-spread diffusion and adoption of emerging technologies across economies and societies, techno-economic paradigms fundamentally alter the structure our socio-institutional frameworks.

In Perez' model, each technological revolution can be characterized further in terms of a specific life cycle, which, as Perez documents, tends to last around half a century. Perez identifies four distinct phases within such a life cycle: an initial period, which is characterized by explosive growth and innovations and new products; a phase of constellation, in which new industries, infrastructures, and technology systems are built out; the full expansion of innovation; and the last phase, which is defined by technological maturity and market saturation.

Perez defines a technological revolution as a set of interrelated radical breakthroughs—that is, singular innovations—that form a constellation of interdependent technologies. A technological revolution, in other words, is a cluster of clusters, or a system of systems of technological innovations. The recent major breakthrough in information technology, for example, formed such a technology

[4] In an article entitled "Measuring Scientific Revolutions," Mark Buchanan asks how we can "identify a "revolution" out of a background of normality. Can there be small revolutions? And could normal periods of science result not from the absence of revolution but from a barrage of small revolutions that together give the appearance of continuous development?" Exploiting the analogy with geophysics and earthquakes—in particular, the seismological finding that there is similarity in the mechanisms driving earthquakes of all magnitudes—he asks whether the evolution of science might be similar. If so, as Buchanan suggests, it might be possible to develop an "intellectual Gutenberg-Richter law": "If one could measure the size of a scientific revolution (perhaps through studies of citations?), and so estimate the magnitude of the rearrangement it entails, then one might similarly quantify the dynamics of scientific understanding. A study of the distribution of the sizes of episodes of scientific change would then reveal whether the Kuhnian dichotomy of normal and revolutionary is really true, or if science instead develops by way of revolutions of all sizes" (Buchanan 1996).

system around microprocessors and other integrated semiconductors, from which new technological trajectories opened up: personal computers, software, telecommunication, and the internet emerged from the initial technological system. These new technological systems subsequently created strong inter-dependence and feedbacks between technologies and markets. The defining features of technological revolutions—as opposed to a random collection of singular innovations—are thus the following: (1) they are interconnected and interdependent in their technologies and markets, and; (2) they have the disruptive potential to radically transform the rest of the economy and society.

Historically, Perez identifies five such major technological meta-systems, which were initially triggered by a technological (or scientific) breakthrough and, then, expanded across industries and economies. The first such disruption of the late 18th century was organized around the mechanization of factories, water power, and the canal networks. This was followed by the second revolution, which initiated the age of steam and railways. In the late 19th century, electricity, steel, and heavy engineering intensified international trade and globalization. In the last century, two technological revolutions transformed our economic and industrial system: the age of oil, mass production and the automobile was followed by the era of information and communication technology.

What made these technological disruptions revolutionary were not only the new interrelated technologies, industries, and infrastructures but their transformative potential defined in terms of extraordinary increases in productivity that they enabled. When a technological revolution propagates across industries and economies, it radically transforms the cost structure of production by providing new powerful inputs (such as steel, oil, or microelectronics). Thereby, it unleashes new innovations and interrelated technological systems, which renew existing industries and create new ones.

Perez provides a powerful framework for understanding the dynamics of technological change. However, as Perez emphasizes, her model applies to a large historical time-frame of around half a century. Consequently, her model of innovation has—as it is the case with most cyclical or wave theories of historical processes—limited power to diagnose, explain, and predict phenomena occurring at smaller time-scales. In other words, while it illuminates patterns of innovation that occur at larger historical time-scales, Perez's conceptual model does not advance our understanding of innovation processes that occur at smaller historical time-scales.

A fundamental assumption guiding this thesis is that—if we want to advance our understanding of the drivers of technological innovation—we need to disaggregate conceptual abstractions, such as "innovation" or "progress," into the key components from which innovation emerges in a bottom-up manner. In other words, we need to take into account the dynamics between different agents, such as entrepreneurs, investors, technologists, or policy-makers, their concrete visions, ideas, or motivations, and the specific socio-historical environment in which they are embedded. What is needed, therefore, is a conceptual model of innovation that can be applied at different levels of analysis. The Social Bubble Hypothesis, which has been developed in our research group and which we mobilize and refine in this thesis, provides precisely the conceptual model of innovation that helps us not only better understand the dynamics of innovation at the micro level, but also its interaction with the emergent patterns that occur at the macro level and which are captured by Perez' framework. The Social Bubble Hypothesis needs, thus, not to be understood as an alternative or competing model of innovation, but one that complements and extends Perez' framework. Whereas Perez' model, for example, identifies and describes the cycles of technological innovation at a larger historical scale, the Social Bubble Hypothesis posits the mechanisms that have given rise to previous waves of innovation, which, in turn, can help us to derive insights into the enabling conditions for future innovations.

1.2 The Social Bubble Hypothesis

At its core, the Social Bubble Hypothesis holds that the dynamics of technological innovation share a deep similarity with financial bubbles. Analogous to speculative manias, social interactions between enthusiastic supporters of an idea, concept, or project trigger a positive feedback cycle, which leads to and reinforces widespread endorsement and extraordinary commitment by those involved in the respective project beyond what could be rationalized by a standard cost–benefit analysis.

Now, the dynamics that govern the formation of financial bubbles are often considered to be negative. The standard view in the economics and finance literature holds that speculative financial bubbles form when unrealistic expectations about future cash flows decouple prices temporally from fundamental valuations. On this view, bubbles are the financial expression of "popular delusions," the "madness of crowds," or "irrational exuberance" (see Mackay 2012; Shiller 2015). Fueled by underlying self-reinforcing feedback-loops of imitation and herding behavior, prices elevate until a crash drives them back to fundamental values (see Sornette 2017). As unrealistic and excessively optimistic expectations about the future fail, bubbles are considered to have primarily societally and economically destructive effects. Indeed, bubbles are persistent and inevitable historical phenomena: from the Dutch tulip mania in 1637 and the South Sea Bubble of 1720, through the great crash of October 1929, Black Monday of 1987, and the financial crisis of 2007, speculative bubbles and market crashes invariably punctuate our history.

While some bubbles have been historically associated with negative outcomes, such as financial crashes and the destabilization of the wider economic system, the formation of certain bubbles can be understood as an important process for innovation in various domains. As some financial bubbles deploy the financial capital necessary to fund disruptive technologies at the frontier of innovation, they are, thus, capable of accelerating breakthroughs in science, technology, and engineering. By generating positive feedback cycles of excessive enthusiasm and investments—which are essential for bootstrapping various social and technological enterprises—bubbles can be societally and economically net-beneficial. In other words, without the innovative spillover-effects of financial bubbles, many technological innovations or large-scale societal projects might never have happened. As shown in earlier case studies of the Human Genome Project and the Apollo Program (see Gisler and Sornette 2009, 2010; Gisler et al. 2011), the complex networks of social interactions between over-enthusiastic supporters and the resulting reduction in collective risk-aversion have catalyzed the formation of many large-scale scientific or technological projects. It is, therefore, not surprising that financial bubbles and speculative manias have been—as documented by Perez—the engine of the technological revolutions that have fundamentally transformed our economic, social, and technological systems over the last three centuries.

In Perez's model, each technological revolution is triggered by a financial bubble, which allocates excessive capital to emerging technologies. She has identified a regular pattern of technology-diffusions. There is an "installation" phase in which a bubble drives the installation of the new technology. This is followed by the collapse of the bubble or a crash, to which Perez refers to as the "turning point." After this transitional phase—which occurred, for example, after the first British railway mania in the 1840s, or, more recently, after the dotcom-bubble in the early 2000s—a second phase is unleashed: the "deployment" phase, which diffuses the new technology across economies, industries, and societies. The economic exhaustion of the technological revolution and excessive financial capital, which searches for new investment opportunities, can, then, give rise to the next technological revolution.

The boom-and-bust pattern, which Perez identifies in large-scale cycles of technological innovation that span several decades, can also be identified in the dynamics driving scientific, engineering, and entrepreneurial projects. Similar to a large-scale technological revolution at the macro-level, a social bubble, which, as mentioned, can form at the micro scale, progresses through the following four idealized phases:

- In the initial phase, there is the invention of a specific idea, project, or technology, which is strongly supported by a small group of participants (such as technologists, investors, entrepreneurs, early adopters, etc.)

- The inflated expectations or "irrational" exuberance around the idea, project, or technology creates a self-reinforcing feedback loop, which reduces risk-aversion and attracts increasing attention and public and/or private investment;

- The investment and attention validate the idea, project, or technology, which, in turn, triggers a capital spending cycle that results in the proliferation of various ventures and accelerated price growth of corresponding firms trading in public and/or private markets;

- There is saturation around the idea, project, or technology, potential termination of the project, or exhaustion of interest and adoption, and, potentially, a decrease of capital inflows, which can lead to a corresponding correction or crash in public and/or private market valuations.

The insight that technological innovation exhibits the dynamics of bubbles is particularly important for understanding the adoption and diffusion of emerging technologies. A key-feature of social bubbles is that the excessive enthusiasm and over-optimistic expectations of those involved increase collective risk-tolerance, which results in the de-risking of a bleeding-edge technology, project, or idea around which the bubble has formed. It is precisely this bubble-induced increase in societal risk-tolerance that stimulates and reinforces excessive investment and interest in a specific idea or technology.

Now, by integrating the micro as well as macro level of analysis, the Social Bubble Hypothesis provides a powerful theoretical framework to understand the dynamics of innovation that circumvents the Scylla of technological determinism and the Charybdis of social constructivism (see Dafoe 2015). Both philosophical views fail to fully capture the dynamics of technological change: Whereas extreme articulations of technological determinism ascribe agency to technology and conceptualize it as a quasi-autonomous process, which, by following an internal logic, controls the historical evolution of our techno-social systems, the extreme version of social constructivism holds that the development of technology is exclusively determined by its social context, human agency, and irreducible historical contingencies. Both views encapsulate valuable insights into the nature of technological innovation. However, as different processes unfold at different scales of analysis, instead of subscribing to either view, they need to be synthesized. The Social Bubble Hypothesis—which is informed by a complex systems-understanding of technological change—fuses the techno-determinist and social-constructivist views of technological progress. It recognizes that, in multi-leveled complex systems, certain recurring patterns, waves, or cycles emerge only at larger scales of analysis, which may be hard to detect on smaller scales of analysis (see Sornette 2008). In the case of technological change, certain deterministic macro-patterns may arise that are not apparent at or explainable at a smaller scale of analysis. In other words, the Social Bubble Hypothesis recognizes simultaneously the methodological validity of analyzing technological innovation at the level of the concrete social context in which technological change unfolds, the interaction of human actors, and irreducible historical contingencies as well as the scale of analysis where technology exhibits emergent trends, patterns, or an internal logic of development, which cannot simply get reduced to social contexts or contingent historical factors.

1.3 Structure of the Thesis

This thesis consists of three self-contained research papers. Each paper is motivated by the question concerning the essence and structure of technological innovation. While providing an exhaustive answer is beyond the scope of this project, each paper aims to contribute to a general theory of technological innovation, which illuminates key-components and drivers of innovation.

This thesis is divided into three parts. The first research paper aims to extract insights into the dynamics of scientific progress on the basis of a historical and conceptual reconstruction of the interaction between physics and economics. The paper analyzes the conceptual and methodological cross-fertilization between these two scientific fields and tracks the emergence of econophysics, a relatively novel field that mobilizes the models and methods of statistical physics to advance our scientific understanding of financial markets. In particular, the paper focuses on the re-conceptualization of financial markets as dynamic complex systems, which has been developed and promoted by proponents of econophysics. The paper shows—on the basis of the contemporary debate around the alleged quantitative intractability of social reflexivity—how scientific advances in physics can dissolve the apparent incommensurability between the methods of social sciences and physics. The paper concludes by positing that—beyond importing methods and models from physics—novel approaches to advance our scientific understanding of financial markets also need to integrate critical insights from the biological sciences.

The second paper advances the Social Bubble Hypothesis. Based on a case study of the invention and development of Bitcoin—the first and most important cryptocurrency protocol—the paper analyzes the social dynamics that govern the adoption and diffusion of this emerging technology. By applying the conceptual framework of the Social Bubble Hypothesis, the paper shows how a hierarchy of repeating and exponentially increasing series of bubbles and hype cycles has catalyzed Bitcoin's adoption. Each bubble, which was fueled by the extraordinary commitment of early adopters, has attracted a new cohort of adopters, which, in turn, has resulted in increasing price accelerations. By dissecting the social dynamics shaping Bitcoin's evolution, the paper further shows how these techno-economic feedback loops, which drive technological adoption, are, in the case of Bitcoin, hard-coded into the protocol itself. In particular, the paper uncovers a deep similarity between the early adopters' excessive belief in and commitment to Bitcoin and religiosity. As the paper shows, this quasi-religious dimension of Bitcoin is an essential feature of its diffusion.

The last paper further extends and refines the Social Bubble Hypothesis by applying it to the clean-tech bubble that formed in the last decade. While it did not coalesce around a particular technology—the bubble formed across a cluster of interrelated technologies, such as solar, biofuels, wind, or batteries—the bubble was coordinated around a shared narrative. The paper analyzes how narratives have historically shaped the image and adoption of energy. In the case of clean or renewable energy technologies, a quasi-religious narrative around "crisis," "catastrophe," and "salvation" has created the emotional resonance that is essential for realizing a bubble's potential for contagion. As the paper shows, the clean-tech bubble is a prototypical social bubble, which progressed through the different phases identified above. The bubble was fueled by a compelling narrative that triggered a cycle of public and private over-investments that funded and parallelized trial-and error experimentation in a plethora of technologies and materials. The paper concludes by contrasting and comparing the clean-tech bubble of the last decade with the contemporary resurgence of a bubble now forming in "sustainability," "climate tech," and "ESG." The paper advances the claim that—even though the first instantiation of the bubble resulted in a myriad of spectacular failures and losses—the clean-tech bubble de-risked a novel category of bleeding-edging technologies and catalyzed a massive reduction in costs, thereby incentivizing the wider adoption of clean technologies a decade after the first bubble burst.

References

Bostrom, N., *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press (2014).

Buchanan, M., Measuring Science Revolutions, *Nature* 384.6607, 325-325 (1996).

Cauwels, P. and D. Sornette, Are 'Flow of Ideas' and 'Research Productivity' in Secular Decline?, *Swiss Finance Institute Research Paper No. 20-90* (2020).

Cowen, T., *The Great Stagnation: How America Ate All the Low-Hanging Fruit of Modern History, Got Sick, and Will (Eventually) Feel Better*, Penguin (2011).

Dafoe, A., On Technological Determinism: A Typology, Scope Conditions, and Mechanism, *Science, Technology, & Human Values* 40.6, 1047-1076 (2015).

Gisler, M. and D. Sornette, Bubbles Everywhere in Human Affairs, chapter in book entitled *Modern RISC-Societies. Towards a New Framework for Societal Evolution*, L. Kajfez Bogataj, K.H. Mueller, I. Svetlik, N. Tos (eds.), Wien, Edition Echoraum 137-153 (2010).

Gisler, M. and D. Sornette, Exuberant Innovations: The Apollo Program, *Society 46*, 55-68 (2009).

Gisler, M., D. Sornette and R. Woodard, Innovation as a Social Bubble: The Example of the Human Genome Project, *Research Policy* 40, 1412-1425 (2011).

Gordon, R., *The Rise and Fall of American Growth*, Princeton University Press (2016).

Hayek, F. A, *The Counter-Revolution of Science*, Free Press (1952).

Hegel, G.W.F., *Phenomenology of Spirit*, Duke University Press (2007).

Heidegger, M., Die Frage nach der Technik, *Vorträge und Aufsätze*, Neske (1990).

Kuhn, T.S., *The Structure of Scientific Revolutions*, University of Chicago Press (2012).

Kurzweil, R., *The Singularity Is Near: When Humans Transcend Biology*, Penguin (2005).

Mackay, C., *Extraordinary Popular Delusions and the Madness of Crowds*, Simon and Schuster (2012).

Marx, K., *Grundrisse: Foundations of the Critique of Political Economy*, Penguin Books (1993).

Meadows, D. and J. Randers, *The Limits to Growth: The 30-year Update*, Routledge (2012).

Perez, C., *Technological Revolutions and Financial Capital,* Edward Elgar Publishing (2003).

Popper, K., *The Poverty of Historicism*, Routledge (2013).

Shiller, R., *Irrational Exuberance*, Princeton University Press (2015).

Sornette, D., Nurturing Breakthroughs; Lessons from Complexity Theory, *Journal of Economic Interaction and Coordination* 3, 165-181 (2008).

Sornette, D., *Why Stock Markets Crash (Critical Events in Complex Financial Systems)*, Princeton University Press, first edition 2002, second edition (2017).

Sornette, D. and P. Cauwels. 1980–2008: The Illusion of the Perpetual Money Machine and What it Bodes for the Future, *Risks* 2.2, 103-131 (2014).

Spengler, O., *The Decline of the West*, Oxford University Press (1991).

Strauss, L., Progress or Return? The Contemporary Crisis in Western Civilization, *Modern Judaism*, 17-45 (1981).

Thiel, P., The End of the Future?, *National Review* 3.10 (2011).

Vinge, V., Technological Singularity, *VISION-21 Symposium sponsored by NASA Lewis Research Center and the Ohio Aerospace Institute* (1993).

2. Science and Progress: The Cross-Fertilization between Finance and Physics

Full Reference:

Huber, T. and Sornette, D. (2016), Can there be a Physics of Financial Markets? Methodological Reflections on Econophysics. *European Physical Journal*, Special Issue, "Can Economics be a Physical Science?", Special Topics 225, 3187-3210.

Abstract

We address the question whether there can be a physical science of financial markets. In particular, we examine the argument that, given the reflexivity of financial markets (i.e., the feedback mechanism between expectations and prices), there is a fundamental difference between social and physical systems, which demands a new scientific method. By providing a selective history of the mutual cross- fertilization between physics and economics, we reflect on the methodological differences of how models and theories get constructed in these fields. We argue that the novel conception of financial markets as complex adaptive systems is one of the most important contributions of econophysics and show that this field of research provides the methods, concepts, and tools to scientifically account for reflexivity. We conclude by arguing that a new science of economic and financial systems should not only be physics-based, but needs to integrate findings from other scientific fields, so that a truly multi-disciplinary complex systems science of financial markets can be built.

Eur. Phys. J. Special Topics **225**, 3187–3210 (2016)

DOI: 10.1140/epjst/e2016-60158-5

THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS

Regular Article

Can there be a physics of financial markets? Methodological reflections on econophysics

Tobias A. Huber[1,a] and Didier Sornette[1,2,b]

[1] ETH Zurich, Department of Management, Technology and Economics, Scheuchzerstrasse 7, CH-8092 Zurich, Switzerland
[2] Swiss Finance Institute, 40, Boulevard du Pont-d'Arve, Case Postale 3 1211 Geneva 4, Switzerland

Received 10 May 2016 / Received in final form 15 September 2016
Published online 22 December 2016

Abstract. We address the question whether there can be a physical science of financial markets. In particular, we examine the argument that, given the reflexivity of financial markets (i.e., the feedback mechanism between expectations and prices), there is a fundamental difference between social and physical systems, which demands a new scientific method. By providing a selective history of the mutual cross-fertilization between physics and economics, we reflect on the methodological differences of how models and theories get constructed in these fields. We argue that the novel conception of financial markets as complex adaptive systems is one of the most important contributions of econophysics and show that this field of research provides the methods, concepts, and tools to scientifically account for reflexivity. We conclude by arguing that a new science of economic and financial systems should not only be physics-based, but needs to integrate findings from other scientific fields, so that a truly multi-disciplinary complex systems science of financial markets can be built.

1 Introduction

Driven by the computerization of financial markets and research technologies, econophysics has been introduced in the early 1990s. Econophysics, as a distinct scientific field of research, emerged as a reaction among physicists and some economists to the failure of standard economic theory to realistically model the complex behavior of financial systems. The early developments of the econophysics literature have been focused on studying and extending a set of so-called "stylized facts," many of them previously identified in financial economics, which are defined as robust empirical features that generalize across markets and asset classes. A rich diversity of econophysical models, such as agent-based, evolutionary and minority-game models and model-driven theories of large price fluctuations, bubbles and market crashes have

[a] e-mail: tobhuber@student.ethz.ch
[b] e-mail: dsornette@ethz.ch

evolved over the last two decades. By illuminating the statistical properties of financial return distributions and their underlying social mechanisms, which can generate systemic instabilities, econophysics has since then significantly advanced our understanding of financial markets.

However, a dominant criticism of the quantitative aspirations to build a physical science of financial markets has emerged. This criticism is encapsulated in the claim that "financial modeling is not the physics of markets," as expressed by Emanuel Derman, a prominent former Wall-Street "quant" and theorist [1]. Another well-known practitioner, the hedge-fund manager and philosopher George Soros, who put forward the concept of market reflexivity [2], argued that financial reflexivity – i.e., the positive feedback between expectations and prices that drive market dynamics – demands a new version of the scientific method as physics cannot handle the complexity of reflexive behavior.[1]

In this paper, we provide a methodological reflection on econophysics, which is guided by the fundamental question whether there can be a physics of markets. This higher-level methodological approach allows us to assess the theoretical contributions of econophysics and identify the main challenges in building a new science of economic or financial systems. In particular, by reflecting on the differences and similarities between physical and social systems, respectively physics and economics, we examine whether econophysics is methodologically equipped to approach the problem of reflexivity.

The paper is organized around three parts. Section 2 gives a brief overview of the mutual cross-fertilization between physics and financial economics. This section highlights how physics and economics methodologically diverged – as indicated by the different uses that data and models have with respect to theorizing. It further illustrates the methodological difference by comparing how physics and financial economics approach the "excess volatility puzzle." In Section 3, we consider how econophysics has contributed to an advanced understanding of economic and financial systems. We present here the ideas of financial complexity and causality that flow from econophysics. We consider the introduction of complexity and heterogeneity to financial economics as one of the most important contributions of the relatively new field of econophysics. By rendering idealized assumptions about representative agents and the equilibrium state of financial markets empirically more realistic, econophysics has been able to account for the heterogeneity of economic agents and the out-of-equilibrium and non-linear dynamics that characterize complex financial systems. In Section 4, we analyze how econophysics has provided a novel perspective on complexity and causality, which allows for bottom-up and top-down causation. We show that the econophysics view of complexity and causality allows one to quantitatively approach reflexivity in financial markets. We conclude by arguing that, while econophysics has made considerable progress in, for example, reproducing many of the stylized facts identified in financial data or elucidating the social or behavioral mechanisms underlying market dynamics, it needs to reach beyond physics and integrate the concepts, methods, and tools from other disciplines, so that a new multidisciplinary complex systems science of financial markets can emerge.

2 Physics vs. finance

The mutual cross-fertilization between physics and economics has a long history starting well before the emergence of econophysics in the mid 1990s. The history and evolution of economics and physics, from the development of classical and neo-classical

[1] The concept of reflexivity has its origins in sociology and philosophy. See [3–8].

economics to econophysics, is punctuated by various collisions between these two fields. In what follows, we provide a selective overview of the historical interaction between physics and economics and finance, before we consider what distinguishes the development of econophysics from other historical cases of inter-fertilizations between these two fields [9–12].

2.1 A brief history of the cross-fertilization between physics and economics

Historically, one of earliest cases of the "physical attraction" [13] of economics was the conceptual influence Isaac Newton's *Philosophiae Naturalis Principia Mathematics* (1687) [14] exerted on Adam Smith's *Inquiry into the Nature and Causes of the Wealth of Nations* (1776) [15]. In particular, the notion of causative force, which was novel at that time, inspired Smith to conceptualize the dynamics of an economic system analogous to Newtonian physics. By the beginning of the 19$^{\text{th}}$ century, the notion of social and economical laws, paralleling physical laws, became deeply entrenched in economics.

In his *Essai philosophique sur les probabilités*, Pierre-Simon Laplace, for example, showed in 1812 that certain social phenomena, which appear random, exhibit law-like behavior and some predictability [16]. Amplifying this scientification of social phenomena, Adolphe Quetelet, who studied birth and death rates, crimes and suicides, even coined the term "Social Physics," which aimed at identifying empirical regularities in the asymptotic Normal distributions observed in social data [17]. (See Galam et al. [18] and Galam [19] for a re-discovery and extension using agent-based models and mapping to physical particle models.)

In the second half of the 19$^{\text{th}}$ century, the economists Alfred Marshall and Francis Edgeworth imported the concept of equilibrium into economics, drawing on research of Clerk Maxwell and Ludwig Boltzman, who described the macro-equilibrium observed in gases as the result of a multitude of collisions of particles. By analogizing economic activity with the interaction of gas particles, Marshall developed the notion of the equilibrium state of the economy, which still forms the core of current mainstream economics (we come back to this below).

Similarly to a thermodynamic description, which uses a mean-field representation that abstracts from the heterogeneity of the particles and their various microstates, equilibrium theory reduces the rich heterogeneity of economic agents to a single representative agent or firm [20]. Interestingly, the idea of equilibrium, which started to heavily dominate economics in the 1950s and now lies at the heart of the Dynamic Stochastic General Equilibrium (DSGE) models that central banks relay on, was struggling to find acceptance by economists, who had an out-of-equilibrium conception of the economy. However, the idea that economies achieve equilibrium states – i.e., that total demand equals total supply or total consumption equals total output and prices are stable – which was the result of a long maturation process, was subsequently pushed to the methodological extreme in neo-classical economics. Whereas equilibrium in physics is a descriptive feature of physical systems, equilibrium has a normative status in economics; the rich and complex out-of-equilibrium and non-linear dynamics of economic systems are tamed and forced to conform to a set of idealized assumptions and theorems, which are couched in equilibrium terms. In other words, rather than describing how economic or financial systems are, (financial) economics strives to prescribe how they should behave [21,22].

Now, not only has economics been fertilized by physics, but also methods and concepts from economics were transferred to physics. An important historical case, in which the flow of concepts and methods between physics and economics has been reversed, is when the economist and philosopher Vilfredo Pareto described in his *Cours d'Economie Politique* (1897) regularities in income distributions in terms of power laws [23]. As they are closely related to the concepts of scale invariance and universality, power laws have been later widely used in physics and the natural sciences to describe the universal statistical signatures of event sizes such as distributions of city sizes, earthquakes, avalanches, landslides, storms, commercial sales, or war sizes. These distributions are not Gaussian, but "fat-tailed" or sub-exponential distributions. One very important insight, which follows from the nature of fat-tailed distributions of event sizes, which power laws describe, is that the probability of the observation of extreme events is not negligible. Whereas Gaussian distributions attribute negligible likelihood to event sizes that are larger than a few standard deviations from the mean, power law distributions provide more weight to these extreme events, which have been documented in various physical systems [24–26]. As these extreme events often control the long-term behavior and organization of complex systems, power laws became very attractive for physicists. However, despite their appeal, it should also be cautioned that many reported power laws distributions turned out to be of limited validity or spurious [27,28].

Another historically important phase in the historical mutual cross-fertilization between economics and physics was Louis Bachelier's attempt to model the apparent random behavior of stock prices in the Paris stock market. In his PhD thesis *Théorie de la spéculation* (1900) [29], Bachelier, who was a student of Poincaré, developed the mathematical theory of diffusion and solved the parabolic diffusion equation five years before Albert Einstein (1905) established the theory of Brownian motion [30]. The modeling of stock prices as stochastic processes – analogue to the random motion of particles suspended in a gas or liquid – constitutes now one of the fundamental pillars of physics and (financial) economics. Research on fluctuation phenomena in statistical physics, on quantum fluctuation processes in quantum field theory, and on financial time-series in finance is now based on the random walk models and their mathematical sibling, the Wiener process.

The Geometric Random Walk model, which was introduced in economics by Osborn [31] and Samuelson [32], uses the exponential of a standard random walk and now constitutes the theoretical core of the most important theoretical constructs in neo-classical finance and economics, such as Markowitz' portfolio theory [33], the Capital-Asset-Pricing model [34], and the Black-Scholes-Merton option-pricing model [35,36]. The theoretical development of these models, which solidified into what has become known as the neo-classical paradigm, resulted in the complete mathematization of economics and finance. Resting on the fundamental assumptions of equilibrium, rational expectations, and utility maximization, the quantitative aspirations of neo-classical finance and economics have been driven by the belief that is possible to model their theoretical foundations on physics. This "physics envy" [37] was already inherent in Paul Samuelson's *Foundation of Economic Analysis* [38], which provided the theoretical framework of what became neo-classical economics, where he adapted the deductive methodology of thermodynamics for economics. Building on these earlier analogies between physics and economics, the economic and finance literature that followed pushed the mathematization of the field to new extremes. Organized around a few parsimonious postulates and theorems, formalized neo-classical finance and economics often devised analytically rigorous models that sacrifice realism for mathematical elegance. Often, these models have generated predictions that are essentially unfalsifiable. With respect to their validity, the models and theories of neo-classical economics, although inspired by physics in the quest for the quantification of economic phenomena, seemed to have methodologically diverged from what has long been considered the "queen of science." In the remainder of this section, we briefly discuss the different uses data and models have with respect to theory-building in physics, respectively to financial economics. We conclude this Section by briefly presenting the evolution of econophysics.

2.2 "Physics Envy"

Although neo-classical economics and finance have been influenced by the methodology of physics of the 19^{th} century and became axiomatized in the second half of the 20^{th} century, there are substantial differences in the way these two disciplines build and test models. While economics privileges models and theoretical principles over empirical confirmation, data come first in physics and models come second [39]. Understanding the methodological difference between physics and (financial) economics, which is manifested in the different ways models are constructed and falsified in these fields, allows one to assess how physics-based approaches can improve scientific models and theories of financial markets.

Whereas economics can be characterized by an axiomatic and non-empirical methodology [40], many physical models are inductively derived from data generated through experiments or simulations.[2] Model construction in economics and finance is often guided by a top-down approach that prioritizes theoretical principles and idealized assumptions, which go into the models. The emphasis on mathematics in neo-classical economics has resulted in highly stylized financial and economic models, which often omit the key characteristics of financial markets. Although many postulates of standard economic theory have been empirically falsified, many economic models, which suffer from their unrealistic assumptions, have not been revised. By contrast, many physical models are generally theoretically more minimal, data-driven, and falsifiable. Furthermore, many physically founded models, which are constructed in econophysics or behavioral finance and are methodologically inspired by statistical physics or biology, are geared towards elucidating the mechanism underlying the phenomenon to be explained or simulated from a collective or many-body perspective. In contrast, economic theory is more interested in how robust individual characteristics of the (representative) decision maker lead to different equilibria.

It is precisely the methodological difference between data-driven and model-based theorizing that allows one to understand the emergence of econophysics and, further, to assess its contributions. In one of the foundational papers of the field, Stanley et al. [41] state that econophysics begins "empirically, with real data that one can analyze in some detail, but without prior models." While it is difficult to agree with Stanley et al.'s implicit assumption of model-free or theory-free observations, econophysics can be nonetheless characterized in large part by the larger emphasis put on data-dependency of its theories and models. However, early econophysics papers were not empirical but more conceptual. Takayasu et al. [42] for example developed the first agent-based model in order to provide an alternative to DSGE models by incorporating agents' heterogeneous characteristics and the role of extended networks. Bouchaud and Sornette [43] were interested in exploring beyond the ideal arbitrage limit in continuous time using Ito calculus to more general situations amenable to the functional integration techniques that have been developed in particle physics and statistical physics.

How the standard approach differs from econophysics can be illustrated by so-called "puzzles" in economic theory. When empirical observations do not conform to the prediction generated by standard models of economics, economists often label these anomalies "puzzles." While the falsification process in physics would dictate to

[2] However, it is important to note that characterizing physics in terms of experimentation is inadequate as some areas of physics, such as string theory, are often experimentally impenetrable. Often, these areas in physics are highly theoretical and difficult if not impossible to confirm empirically, at least in the foreseeable future. Furthermore, over the past decades, economics witnessed the emergence of empirical approaches, such as experimental or behavioral economics.

reject the model on empirical grounds, the theoretical edifices of neo-classical economics are getting improved so as to account for these anomalous economic phenomena. In other words, data must conform to the normative nature of the mathematically parsimonious and elegant models.[3] One of the most famous anomalies in financial economics is the so-called "excess volatility puzzle," which Shiller [45,46] and LeRoy and Porter [47] have unearthed in the 1980s. It refers to the empirical observation that prices fluctuate too much when compared to their fundamental valuation.

It follows from the efficient market hypothesis (EMH) – which states that the trajectory of stock prices instantaneously incorporates all available information – that the observed price $p(t)$ of a share or index is equal to the mathematical expectation, conditional on all available information, of the present value $p^*(t)$ of actual subsequent dividends accruing to that share. As this fundamental value $p^*(t)$ is not known, it has to be forecasted. For EMH, the observed price is equal to the forecasted price, which can be written as

$$p(t) = \mathrm{E_t}[\mathrm{p^*t}], \tag{1}$$

where $\mathrm{E_t}$ refers to the mathematical expectation conditional on all public information available at time t. As EMH asserts, only new information on the fundamental value $p^*(t)$, which was not available at the time of the forecast, might result in surprising movements in the stock price. If follows from EMH that

$$p^*(t) = p(t) + \varepsilon(t), \tag{2}$$

where $\varepsilon(t)$is a forecast error, which must be uncorrelated with any information at time t as the forecast would otherwise not be optimal, respectively the market not efficient. However, empirical observations of price behavior show that the volatility of the realized price $p(t)$ is much larger than the volatility of the fundamental price $p^*(t)$. This gives rise to the "excess volatility puzzle" as the empirical observation conflicts with the theoretical predictions of the model (2), since, mathematically, the volatility of $p^*(t)$ obtained as the sum (2) cannot be smaller than the volatility of one of its constituents $p(t)$, given that $\varepsilon(t)$ is uncorrelated with $p(t)$.

However, when viewed through the lenses of the logic of physics, there is no "excess volatility puzzle." As physics operates with the concept of causality, the observed price $p(t)$ can be understood as following from fundamentals. In other words, $p(t)$ should be an approximation of $p^*(t)$. Therefore, expression (2) should be replaced by

$$p(t) = p^*(t) + \varepsilon'(t), \tag{3}$$

and there is no longer a volatility excess paradox. The observed realized volatility of $p(t)$, which is larger than the volatility of $p^*(t)$, provides information on the price-formation process, which is not optimal. The introduction of causality shows that the observed price approximates the fundamental price, up to an error of appreciation of the market. It follows from this that price moves have other causes than fundamental valuations: there exists a noise element in the pricing process, which results in the deviation of the observed price from its fundamentals. However, instead of exploring the mechanisms underlying price fluctuations, most economists do not reject the a priori assumptions of standard economic theory. The difference between (2) and (3) captures this fundamental difference in modeling strategies in economics versus

[3] MacKenzie [44] shows how financial models changed reality. For example, the practical implementation of the Black-Scholes-Merton model by traders at the Chicago Board Options Exchange rendered the model more realistic – until the crash of Oct. 1987 broke its validity, requiring the introduction of the so-called "volatility smile" as a fudge to make it continue working.

physics. In its top-down approach, standard economic theory, such as EMH or rational expectations, dictates that economic reality has to conform to the models because the market is supposed to have had time to absorb all information and converge to its optimal representation, the market price. This economic view assumes that all nonlinearities and reflexive loops of infinite order are taken into account [48]. This is the limit of perfect rationality and infinite effective computing power. In contrast, the bottom-up physics-based approach is more myopic or time-dependent, examining the underlying microscopic system in order to illuminate its aggregate or macroscopic properties through a progressive construction process, which could be called "inductive" [49]. It is not assumed that the investors have digested all available information and have developed optimal strategies accounting for those of their peers in infinite loops of reflexivity.

As the example of the excess volatility puzzle above shows, the bottom-up approach of physics, which introduces the concept of causality and imperfect markets, can substantially advance our understanding of economic phenomena or systems. It is precisely the observation that there are parallels between economic and physical systems and that physics can contribute to their understanding that resulted in the emergence of econophysics as a new and distinct field in the 1990s. We briefly discuss in the next section the history of econophysics and what distinguishes it from other historical cases of cross-fertilizations between physics and economics, before we examine in more detail what we consider the most important theoretical contribution of econophysics.

2.3 The emergence of econophysics

The term econophysics – a synthesis of economics and physics – already encodes the theoretical program of the field. Similar to "astrophysics," "geophysics," or "biophysics", econophysics strives to model, predict, and explain economic phenomena by applying tools and concepts from statistical and theoretical physics [50,51]. In this section, we briefly identify the factors that have driven the evolution of econophysics.

Fuelled by the simultaneous computerization of financial markets and academic research, the availability of high-frequency data made the application of physical methods and concepts specifically attractive. One of the most important insights flowing from the last two decades of research in econophysics is that financial markets and their dynamics can be understood as complex adaptive systems, from which a variety of stylized facts emerge. Stylized facts – which represent robust and universal properties that can be identified across data sets from different markets and asset classes – can be understood as emergent properties, which result from the complex and nonlinear interactions of the system's components. Econophysics emerged from attempts to describe these stylized facts, such as volatility clustering, the heavy-tailed nature of return distributions, or the absence of linear correlations between returns [52–54], in terms of statistical frameworks used in physics. This statistical physics approach – which started in the 1960 with Mandelbrot [55–57], Fama [58,59] and Samuelson's [60] re-interpretation of modern portfolio theory using Paretian power laws and stable Lévy distributions [13] [4] – consolidated in the 1990s into a new field, which promised to statistically explain these universal patterns in financial data, which standard economic theory fails to account for. The leptokurticity of financial distributions, which was already unveiled by research in the 1960s following Mandelbrot's foray in the field and subsequent research by Fama himself, Cootner [62] and others, collides

[4] A modern approach generalising Samuelson to the case of distributions of returns with power law tails (not necessary in the Lévy law domain of attraction) was offered by Bouchaud et al. [61].

with standard economic theory, which is methodologically underpinned by Gaussian assumptions. The fundamental pillars of finance theory, such as modern portfolio theory, the Black-Scholes-Merton model of option pricing, CAPM, or value-at-risk, are characterizing the distributions of financial returns within a Gaussian framework. Because of its fundamental assumptions about a Gaussian world where markets are efficient and in equilibrium and economic agents perfectly rational, standard economic theory has invariably failed to account for extreme events in financial systems [63]. Econophysics' promise to model, explain, and predict the non-Gaussian nature of financial systems, their non-linearity and out-of-equilibrium dynamics, accelerated its development during the 1990s.

However, it is important to note that the evolution of econophysics is radically different from the historical cases where physics was infused into (financial) economics. Whereas in the cases we discussed in the previous section, the physical concepts, methods, and tools have been translated into economics – i.e., the physics concepts have been rendered economically meaningful – and integrated into the theoretical framework, econophysics is a new approach to economic or financial systems, which does not incorporate the theoretical foundations of standard economic theory. Econophysics has not simply integrated concepts and methods of statistical physics into the framework of financial economics, but directly applied physics, its concepts and methodologies to economic phenomena, often without building on the standard (neo-classical) theories of financial economics [64]. In other words, econophysics is simply an extension of the physical sciences, which has been motivated by the phenomenological and conceptual similarity between physical and financial systems. Consequently, the fact that econophysicists often filter out standard economic theory, a fact that also concerns the sociological dynamics of these scientific fields, has generated some controversies. While some econophysicists have expressed the desire to replace the theoretical edifice of neo-classical economics with econophysics, some economists claimed that econophysists's ignorance of standard economics theory resulted in a replication of scientific results, which have already been well established in economics [65–67]. It has been argued, for example, that complexity approaches, such as econophysics, are "justifying themselves by how they correspond with already-observed facts, rather than by the new insights they provide" [68]. However, the fact that research in econophysics often strives to identify and explain stylized facts in financial data by reproducing the phenomenon to be explained with computer simulations should be considered a positive contribution. Simulations almost never appear in the standard economic literature[5] as the often non-linear out-of-equilibrium and heterogeneous nature of the phenomena to be modeled cannot be analytically solved by standard economic modeling techniques, which follow the neo-classical dictum of simplicity, tractability and conformity to theoretical principles such as rational expectations with well-defined utility functions. By contrast, simulations have become deeply rooted in the practice of physicists working on economic or financial systems as they allow them to simulate algorithmically the behavior of the interacting components of the system under study. In other words, simulations provide a way to "increase the range of phenomena that are epistemically accessible to us" [72] and which the closed form solutions, demanded by standard economic techniques, fail to model [32]. Whereas standard economic models are in most cases not concerned with underlying causal mechanism emphasizing the transition from the micro-level to the macro-level, adopting the simulation methods used in physics enables econophysicists to build explanatory models of the target system that clarify the self-organization processes at work. By reproducing the basic mechanisms underlying stylized facts, econophysics is able to provide explanations of the various macroscopic patterns and regularities observed in financial data. One can

[5] A notable counter-example is e.g. Schelling's model of segregation, from micro-motives to macro-behaviors [69–71].

consider these generative explanations, which provide micro-specifications of macroscopic patterns, as one of the unique features of econophysics [73].

Now, it is remarkable that econophysics is often methodologically defined. As a cursory scan of some definitions given in relevant papers and textbooks reveals, characterizations of econophysics often appeal to its techniques and methods (such as statistical mechanics, power laws, scale invariance, etc.), rather than its theoretical focus. For example, Mantegna and Stanley [51] write that the "word econophysics describes the present attempts of a number of physicists to model financial and economic systems using paradigms and tools borrowed from theoretical and statistical physics" and the "characteristic difference [from the standard economic approach – TH/DS] is the emphasis that physicists put on the empirical analysis of economic data" (ibid.); Stanley et al. [41] state that the econophysics advances "in the spirit of experimental physics"; and Burda et al. [74] define econophysics as a "a quantitative approach using ideas, models, conceptual and computational methods of statistical physics applied to economic and financial phenomena. "For our part, what we consider the most central contribution of econophysics is that – as Stanley et al. [41] write in their foundational paper – "economic systems are treated as complex systems." Econophysics, in other words, provides a fundamental re-conceptualization of economic systems. The defining feature of econophysics is thus less that it applies the techniques of statistical mechanics to financial systems, but primarily the insight that the dynamics of financial systems are best understood as emergent properties of a complex adaptive system. This complexity approach to financial market is what renders econophysics fundamentally different from standard economic theory. In fact, its roots go back to the complexity approach applied to economics that was promoted by the Santa Fe Institute created in 1984 under the particular push of the Nobel economist K. Arrow together with two Nobel physicists P.W. Anderson and M. Gellman [75]. The Santa Fe Institute has been pushing further the unifying concept of "complex adaptive systems" [76,77], with a strong anchor in biological ecologies and evolutionary selection. In contrast, econophysics is a direct descendant of the more traditional statistical physics and experimental physics approach.

In the next section, we analyze the concept of complexity in econophysics in more detail and explore what novel insights about the dynamics of financial markets flow from it. In particular, we look at how econophysical complexity provides a new understanding of causality, before we conclude by examining whether econophysics can methodologically account for the reflexive behavior that financial markets exhibit.

3 Complexity, causality, and reflexivity in econophysics

Viewing physical, biological, or social systems through the lens of complexity has revolutionized many scientific fields over the last decades. The emerging new complexity paradigm, which affected fields as diverse as physics, biology, ecology, or sociology, has furnished explanations to phenomena such as symmetry-breaking, dis-equilibrium, or spontaneous instabilities. Collective or macroscopic phenomena – such as these "critical" phenomena that conflict with the symmetry and equilibrium paradigms, which have dominated physics and biology – are in a bottom-up way understood to be generated by the microscopic interactions of the component parts of a complex system. The aggregate or global states of the system, however, are emergent properties and are not reducible to a particular configuration of the constituents. What is particularly interesting is that these macro-patterns or phenomena can be realized by different systems, i.e., they sometimes exhibit, what physicists call, "universality." Intuitively, economic or financial systems are natural candidates for a complexity approach. In particular, neo-classical economics with its emphasis on equilibrium states, homogeneity of economic agents, and Gaussian distributions, seems especially ripe to transform via the complexity treatment. In the next section, we define a few conceptual properties of complexity and consider briefly how they relate to financial systems, before we examine in more depth how econophysics approaches complexity methodologically.

3.1 Complexity

The study of out-of-equilibrium dynamics (e.g. dynamical phase transitions) and of heterogeneous systems (e.g. glasses, rocks) has progressively made popular in physics and then in its sisters branches (geology, biology, etc.) the concept of complex systems and the importance of systemic approaches: systems with a large number of mutually interacting parts, often open to their environment, self-organize their internal structure and their dynamics with novel and sometimes surprising macroscopic "emergent" properties. The complex system approach, which involves seeing interconnections and relationships, i.e., the whole picture as well as the component parts, is nowadays pervasive in the control of engineering devices and business management. It is also plays an increasing role in most of the scientific disciplines, including biology (biological networks, ecology, evolution, origin of life, immunology, neurobiology, molecular biology, etc.), geology (plate-tectonics, earthquakes and volcanoes, erosion and landscapes, climate and weather, environment, etc.), economics and social sciences (including cognition, distributed learning, interacting agents, etc.). There is a growing recognition that progress in most of these disciplines, in many of the pressing issues for our future welfare as well as for the management of our everyday life, will need such a systemic complex system and multidisciplinary approach.

A central property of a complex system is the possible occurrence of coherent large-scale collective behaviors with a very rich structure, resulting from the repeated non-linear interactions among its constituents: the whole turns out to be much more than the sum of its parts. Recent developments suggest that non-traditional approaches, based on the concepts and methods of statistical and nonlinear physics coupled with ideas and tools from computation intelligence, could provide novel methods into complexity to direct the numerical resolution of more realistic models and the identification of relevant signatures of impending catastrophes. In the following, we identify a few key properties of complex systems that are particularly interesting when one considers financial systems. Most definitions of complexity follow more or less Philip Anderson's classic formulation of complexity, which he gives in his seminal paper "More Is Different" [78]:

> The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles. Instead, at each level of complexity entirely new properties appear [...].

Financial systems can be considered to qualify as complex systems as they are composed of different levels of complexity that generate new and emergent properties. Hormonally-induced changes in the endocrine system influence, for example, financial risk taking and decision making of individual traders [79] whereas traders are socially affected by imitation and herding dynamics [80]. The resulting global dynamics of markets can in turn trigger collective trader behavior, which can cause financial market crashes [81,82]. Each of these examples can be considered as an emergent property of the underlying level of complexity. Financial systems can thus be characterized as containing a large number of interdependent and mutually interacting microscopic sub-units, which produce non-linear and stochastic dynamics from which novel macroscopic states or properties emerge. The following four properties form the conceptual kernel of complexity relevant for econophysics [82,83]:

- *Non-Linearity*: Whereas in linear models the output is proportional to the cause, non-linearity complicates the relation between output and cause [84,85]. Hence, linear extrapolations from microscopic properties to macroscopic phenomena in non-linear systems are bound to fail. Furthermore, complex non-linear systems exhibit positive and negative feedback [86]. Negative feedback result when the fluctuations in the output of a system tend to be reduced compared to the disturbances or changes in the input. Under negative feedbacks the system is stable and tends to reverse to the mean. In contrast, positive feedback occurs when a small change in the input of a system or a disturbance of its parameter amplifies into large system-wide perturbations. While negative feedback tends to stabilize the system by forcing it back into its equilibrium state, positive feedback tend to destabilize the system and results in strongly nonlinear oscillations, out-of-equilibrium dynamics, chaotic behavior, or even to transient singular dynamics associated with changes of regimes. In the financial context, positive feedback is interesting as higher (lower) prices, for example, feed back into trader's behavior, thereby accelerating the upward (downward) price trajectory [80] into characteristic transient super-exponential trajectories [87–89].
- *Emergence*: Emergence is a notoriously slippery concept. Minimally, macroscopic properties of a complex system can be characterized as emergent when they are irreducible to the mutually interacting microscopic parts, which generate novel global phenomena on a higher level of complexity. In other words, the aggregate or global phenomena or patterns transcend the dynamics, properties, and configurations of the system's sub-units. Concretely, the knowledge of the microscopic laws does not predict the macro-behavior as many different microscopic laws can lead to the same large-scale behavior while some apparently innocuous changes actually lead to revolutionary alterations at the global level. In that sense, financial markets can exhibit emergent behavior, which is not shared by its constituents [90]. Before crashes, statistical signatures can be identified that result from the increasing global cooperativity and self-organization of markets. A super-exponential accelerating price decorated by log-periodic oscillations reflecting large-scale volatility organization indicates that the market as whole anticipates crashes before its individual parts do [73,91–95].
- *Criticality*: Criticality is another key characteristic of complex systems that can be derived from the microscopic organization and long-range dependence of the system elements. A complex system exhibits criticality when local influences propagate over long distances and it becomes exceedingly sensitive to small perturbations, which can cause massive changes in the overall behavior. A number of extreme events, such as market crashes, have been proposed to belong to the class of critical phenomena. In the case of financial markets, this criticality, which can be observed in the finite-time singular behavior of financial prices before crashes [95], results from the high correlations and cooperation between the system's elements, such as traders, banks, etc. When the market matures towards a so-called "critical point" leading to an unstable phase, a small distortion might trigger a price collapse [86,91,92]. Applying and generalizing the physics of critical phenomena to financial systems has deepened our understanding of market crashes and speculative bubbles, which cannot be explained from the perspective of standard economic theory. When financial systems undergo critical phase transitions, the origin cannot be traced to an exogenous source such as arrival of major news – as post-mortem analyses have revealed – but the critical state often arises endogenously.

– *Qualitative Universality* [96]: The observation that complex systems, under certain conditions, exhibit universality was one of the factors that encouraged physicists in their belief that econophysics could develop as a distinct scientific field. Roughly speaking, universality refers to the fact that physical, biological and social systems have similar properties that can be generalized across many different system classes. Properties of a system are considered universal when they can be described independently of the details of the microscopic organization of its sub-units. While many descriptions of physical systems rely on scale-dependent parameters, scale-invariant descriptions of its behavior – which are often governed by power laws – begin to dominate when complex systems enter a critical phase transition. A specialized and narrower definition of universality is the concept of self-similarity, which in critical phenomena is associated with the notion of universality classes characterized by identical exponents (or fractal dimensions). Studies on herding behavior amongst traders have suggested that financial systems, when they approach a critical point, are self-similar across scales [91,92].

Taken together, these characteristics form the core of the concept of complexity around which most research in econophysics is organized. In some studies, these features are explicitly defined, in others they are part of the theoretical background that is assumed. What is important now is that these definitions of complexity entail a novel conception of causality. This is very relevant as it allows one to tackle the question whether econophysics is equipped with the methods and concepts required to deal with the problem of reflexivity, which is endemic in social systems.

3.2 Causality

From the characterization of complex systems in terms of emergence follows that emergent properties are irreducible to or not derivable from the lower levels of complexity. When a complex system exhibits emergence, the higher macro-level, which is the locus of emergent properties, exerts causal influence on the lower micro-level substrate from which these novel features of the system have emerged. This causal propagation of effects from higher to lower levels of complexity is often referred to as "*downward causation*" or "macro-determination" – a term that originated in the context of complex biological systems [97,98]. In physics, it is called "direct cascade" in analogy with the cascade of energy from large to small scales in hydrodynamic turbulence [99]. The existence of such a downward cascade, which collides with the more standard micro to macro cascade, is the main reason for the lack of a generic theory of complex systems. The renormalization group theory has solved essentially the problem of the micro to macro cascade for a restricted class of statistical physics systems [100]. But the occurrence of top-down influences and its interactions with the bottom-up micro-macro cascade make system dynamics and organization much richer and elaborate than we can presently fully fathom. Concretely, financial systems can be considered to exhibit downward causation as the aggregate dynamics of markets can causally influence the microscopic behavior of its individual sub-units [101]. For example, phases of extreme market volatility induce a reaction by traders, which often results in more volatile price behavior [102]. Markets can thus be described through a macro-structure that causally affects all its micro sub-units.

In the context of financial systems, econophysics provides a useful theoretical framework to model, explain, and predict the emerging statistical properties of the aggregate level, which do not exist on the microscopic level. By contrast, standard economic theory is epistemologically and methodologically simply not equipped to fully handle the phenomenon of downward causation. The methodological reason for this failure is the conceptual reductionism that is deeply engrained into standard economic theory. Many current macro-models in standard economics are built on the assumption that the macro is fully reducible to the micro. This reductionist approach in standard economic theory centers on the representative agent framework, which gets rid off the heterogeneity of economic agents [64,103]. In the standard economic literature, there "are no assumptions at the aggregate level, which cannot be justified, by the usual individualistic assumptions. This problem is usually avoided in the macroeconomic literature by assuming that the economy behaves like an individual" [104]. As a consequence of this conceptual reductionism, the methodology of neo-classical economics and finance cannot model the connection between the micro-level with the macro-level beyond the assumption that, roughly speaking, the macro level obeys the same laws as the micro level. By equating the micro-level, which is populated with diverse and heterogeneous economic agents and shaped by their mutual non-linear interactions, to the macro-level, the standard economic framework produces highly unrealistic models, which cannot capture the complex dynamics and evolution of financial markets. In particular, the standard economic approach is ill-suited to explain the occurrence of crises and systemic risks. In the build-up towards the great 2008 crisis in the USA, this blindness was embodied by the claim that the economy has transitioned into a new greater level of functioning, dubbed the "Great Moderation", while in reality this apparently improved performance was bought at the cost of non-sustainable debt and financialization and the build-up of a virtual world increasingly disconnected from the real economic world [105].

The epistemological reason for this breakdown of the standard economic model has to do with the "positive" epistemology of economics, which Milton Friedman introduced in 1953 and that still dominates neo-classical economics and finance today. For Friedman, the ultimate criterion for the scientific validity of a scientific theory lies in the quality of its predictions and not in the realism of its assumptions. Whereas a prediction of a model or hypothesis can be falsified, the underlying theoretical edifice is immune to falsification [106]. It is precisely this conception of a positive scientific methodology that gave rise to the axiomatic and unfalsifiable paradigm of neo-classical economics and finance to which the formation of econophysics reacted.

In the next section, we examine how econophysics approaches and models financial complexity. We then address the more fundamental issue of whether econophysics can methodologically account for reflexivity. The problem of reflexivity in financial markets, in turn, depends on how econophysics can link the micro-level and the macro-level that characterize complex systems. We conclude by discussing briefly the problem of reflexivity and how it relates to the econophysics conception of financial markets as complex systems.

3.3 Reflexivity

Given that complex systems, which can be physical, biological or social in nature, are characterized in terms of spontaneous dynamics and properties that emerge from the non-linear interactions and interdependencies of the system's sub-units, a fundamental question arises: what is the source that causes complex systems to behave in this way?

In the context of financial systems, we have seen with the example of the "excess volatility puzzle" that the behavior of asset prices cannot simply be explained in terms of changes in their fundamental values. There is a noise component in the pricing mechanism associated with bounded rationality and out-of-equilibrium processes, which the efficient market hypothesis fails to account for or, rather, hypothesizes away by assuming perfect collective rationality. The question then becomes what is the source of this noise in financial markets. The dynamics of price series cannot be exclusively explained by external "forces" such as the arrival of news, as massive price changes often occur without the presence of any significant piece of information [107,108]. This means that the evolution of financial markets is also subject to endogenous influences, which originate from within the system. Self-organized criticality, and more generally, complex system theory contend that systems with threshold dynamics that are out-of-equilibrium slowly relax through a hierarchy of avalanches of all sizes. Accordingly, extreme events are seen to be endogenous. In economics, endogeneity versus exogeneity has been hotly debated for decades. A prominent example is the theory of Schumpeter on the importance of technological discontinuities in economic history. Schumpeter argued [109] that "evolution is lopsided, discontinuous, disharmonious by nature [...] studded with violent outbursts and catastrophes [...] more like a series of explosions than a gentle, though incessant, transformation." Endogeneity versus exogeneity is also paramount in economic growth theory.

A useful explanation of the endogenous dynamics of markets is rooted in the concept of reflexivity. Popularized by hedge-fund manager and philosopher George Soros, reflexivity captures the fact that, in social or economic systems, expectations of participants influence the evolution of the system, which, in turns, affects the behavior of participants again [2,8]. Soros writes that "the participants' view influence but do not determine the course of events, and the course of events influences but does not determine the participant's view" [8]. The various positive and negative feedbacks, which can be identified in financial system, reflects this financial reflexivity. Reflexive phenomena can thus be characterized by the collision between downward causality and bottom-up aggregation, as discussed above. The microscopic interactions of elements at the lower level (market participants such as traders, hedge funds, regulators, etc.) generate a new macroscopic level of complexity, which, in turn, changes the dynamics and organization of the lower level.

Sharing a widely held view [1,37,110], Soros argues that the fact that financial systems exhibit reflexivity, i.e., that a substantial part of market behavior has a reflexive or endogenous source, demands a scientific method, which is distinct from the tradition of physics. It is argued that, because social systems – the class to which financial markets belong – are fundamentally different from physical system, they are not susceptible to methods borrowed from statistical or theoretical physics. Furthermore, Soros asserts that standard economic theory is "an axiomatic system based on deductive logic, not empirical evidence" [8] and, consequently, it fails to account for the reflexive behavior of financial systems. Based on the argument that one cannot, analogously to physics, identify invariable and universal laws in social systems, Soros then criticizes the "ill-fated attempt by economists to slavishly imitate physics" [8]. As we have shown, the standard economic models, which rest on the assumptions of equilibrium, perfect rationality, and efficient markets, indeed fail to account for the rich reflexive dynamics that exist in financial markets. However, we assert that the problem lies less in the enslavement of financial economics by physics, but in the choices of the physical and mathematical methods and concepts that are getting transferred to economics. As the discussion of the evolution of neo-classical (financial) economics above has shown, economists have been envying or imitating a kind of physics, which was not sufficient to deal with the complexity of financial markets. It was precisely the methodological and conceptual adoption of statistical and out-of-equilibrium physics methods that opened up the possibility to detect the scale-invariant and universal statistical regularities in financial systems that are fundamentally shaped by social and technological forces.

In order to assess whether econophysics has the potential to enlighten the reflexive nature of financial systems, we dissect in the next section how econophysics approaches the complexity of financial markets and address the question whether reflexivity can be quantified.

4 Modeling complex financial systems

An understanding of complex systems and of their non-linear dynamics and emergent properties demands the comprehensive modeling of the link between the micro- and macro-levels. However, we can identify in the econophysics literature two methodological currents that deal with these different levels of complexity: agent-based and statistical econophysics. Whereas agent-based approaches aim at modeling the micro-level of financial systems, i.e., the interaction between economic agents, statistical approaches in econophysics are concerned with the macro-level, i.e., the aggregate patterns and phenomena generated by the system's sub-units. These different modeling strategies, which Schinckus [111,112] has identified, are, for example, reflected in two important survey papers by Chakraborti et al. [53,113], which are organized around a review of econophysics research dealing with stylized facts and a survey of research concerning agent-based models. In this section, we clarify the different methodological motivations in modeling complexity in econophysics. We then conclude by evaluating whether econophysics can attack the problem of reflexivity.

4.1 Agent-based vs. statistical models in econophysics

The two different strategies of agent-based and statistical modeling approaches in econophysics derive their theoretical foundations both from statistical mechanics and theoretical physics. Agent-based models, which strive to model the microscopic dimension of financial systems, and statistical models, which aim at explaining the macro-regularities or stylized facts that can be identified in financial data, are both data-driven and involve theoretically minimal assumptions. What distinguishes these modeling approaches, however, is their different emphasis on the nature and behavior of economic agents.

Whereas agent-based approaches aim at integrating the learning and adaptive features of market participants, statistical modeling often subdues the individual characteristics of economic agents under the emergent collective organisation. Statistical models often extract stylized facts from past financial time-series by using vast amounts of high-frequency data on prices, volumes, and transactions. The statistical descriptions of these empirical regularities often do not require the specification of the underlying behavioral mechanisms. Obviously, the methodological distinction between the two modeling approaches is not as clear-cut as we present it here. However, in the agent-based literature, we can find research that uses the "order ≡ particle" analogy [113]. Inspired by reaction-diffusion models in physics, Bak et al. [114], for example, simulate price variations on the basis of crowd behavior. The price dynamics, represented as market orders, is mapped onto a model of diffusing and annihilating particles. While this early model has many unrealistic features, the simplicity of the particle-representation inspired richer models based on detailed high-frequency data analyses [115–117]. Often, however, these order models assume so-called "zero-intelligent agents" [118], i.e., economic agents are modeled as particles without any behavioral features. These "particles" obey statistical properties and generate the stylized facts, which are known to exist in financial data, but, they do not have the faculty of anticipation and of forming expectations. When it comes to realism, this atomization of financial dynamics into collisions of unthinking particles can be a drawback of some statistical econophysics models [112]. But it can also be an extraordinary powerful approach for the characterization of the changing risk profiles of financial markets and for predictions [116,117]. This zero-intelligence agent-based model approach thus challenges researchers to identify where higher levels of intelligence might impact the observed structure of financial markets.

Agent-based techniques often compensate for this by enriching their models with adapting, learning, and evolving agents. Similar to the formation of molecules or crystals, these models strive to explain the macro-structures as emerging from the microscopic behavior of a system's heterogeneous components. This strategy results in models of financial markets as adaptive complex systems that evolve in time. Endowing the systems' components with behavioral traits allows for more realistic depictions of market dynamics. However, as LeBaron notes [119], agent-based models, which explain price actions in terms of simple behavioral rules that govern the behavior of market participants, are themselves often exceedingly complex and it can be difficult to isolate the factors responsible for generating the stylized facts. A set of models, which are more faithful in their descriptions of the behavior of real traders and markets, has, for example, introduced behavioral switching mechanisms. These models often divide the population of market participants into two groups – "fundamentalists" and "chartist" or "noise" traders – in order to explain market instabilities, which can be observed in real markets. As these traders can switch between different strategies and states, many empirically observed phenomena such as volatility clustering or herding behavior can be realistically reproduced with this class of agent-based models [120–122]. While statistical models of financial data generate empirically adequate descriptions, this modeling approach needs to be complemented with agent-based modeling. Zhou and Sornette [123] provide, for example, an Ising-model of agent's opinions and how they react to external news. By incorporating behaviorally realistic assumptions, which correspond to evidence in neurobiology and behavioral finance, the model is able to reproduce certain stylized facts that result from crowd behavior (see also the extension of Harras and Sornette [124] to account for the spontaneous emergence of bubbles from over-learning by agents of random news). Put more generally, agent-based models provide micro-foundations to the emergent statistical macro-properties of markets and are able to integrate realistically the heterogeneous features of economic agents, such as the range of preferences, deviations from rationality, and social dynamics such as herding or imitation, which give raise to the stylized facts that statistical approaches seek to reproduce. However, the use of agent-based models is still limited by the difficulties associated with their calibration to empirical data [22].

Given the characterization of reflexivity provided above, it follows that any successful attempt to model the reflexive behavior of market participants must span the micro- as well as the macro-domains of markets. Modeling strategies that only target one level at the expense of the other, seem to inhibit a deeper scientific understanding of reflexivity. Furthermore, it does not do justice to the complexity of financial markets to model solely the connection between their micro- and macro-dimensions. Ultimately, complex systems such as markets are embedded in a wider context or environment; they are part of a nested hierarchy of complexity [125]. In other words, models of reflexivity need to capture additionally the entanglement between exogenous and endogenous causes that influence the global behavior of the system [126]. We now present recent attempts to model reflexivity quantitatively, before we conclude the paper.

4.2 Quantifying reflexivity? Endogeneity vs. exogeneity

As it is evident from the discussion above, it is very difficult to scientifically track complex systems, which involve a multitude of mutually and non-linearly interacting parts from which new and surprising phenomena emerge. Complex systems, therefore, demand a new multidisciplinary approach. Nonetheless, over the past two decades, considerable scientific progress has been made by approaching Nature from a complex system science perspective. The complexity of these physical, biological, and social systems has challenged the previous reductionist approach, which consists of decomposing the system into component parts, such that the detailed understanding of the sub-units was believed to generate understanding of the system itself. By contrast, complex systems science approaches the phenomenon through the interactions and links between the sub-units and the system, thereby accounting for positive and negative feedback and downward causation.

Given the complexity of social systems, it has been argued that it is hard to bring them under scientific control [1]. This is due to the fact that the laws or regularities that govern their behavior are difficult to extract. It is also impossible to isolate social and economic systems and to experimentally manipulate them. In particular, the phenomenon of reflexivity often seems "difficult to identify and impossible to quantify" [8]. Ultimately, however, the quantification of the dynamics of financial markets, including their reflexive behavior, is necessary if we want to advance our scientific understanding of financial systems. Only quantification, which allows for the prediction and control of systems, generates a scientific understanding of the target phenomenon. A new science of financial systems, if it were to be effective, needs to be able to quantify the reflexive dynamics that are intrinsic to markets. Before we conclude, we show that is possible to quantitatively disentangle the phenomenon of financial reflexivity by way of the distinction between endogenous and exogenous factors, which contribute to the dynamics of financial markets.

As already mentioned above, the Efficient Market Hypothesis (EMH) assumes that markets almost instantaneously incorporate the flow of information and faithfully reflects it in prices. The "excess volatility puzzle," amongst other empirical findings, contradicts the main tenet of one of the pillars of standard economic theory. The EMH asserts that, normally, the market efficiently absorbs exogenous shocks and converges towards an equilibrium price, while endogenous processes are absorbed into the price-formation process and disappear as part of the digestion of the exogenous information. Consequently, markets, following the EMH, are only driven by exogenous inputs, and not by endogenous dynamics. However, the reality of financial markets is radically different as price volatility is too high as could be justified by shifts in the underlying fundamental valuations. Furthermore, a variety of studies have refuted the EMH's assumptions that extreme events in financial systems are induced by the exogenous negative impact of information [80,107,116,117]. Similarly to other complex systems, such as fluctuations in turbulent flows, avalanche dynamics, or earthquakes, financial markets exhibit an endogenous dynamic that is very complex, whereas exogenous forces, which drive the system, are often regular and steady [126,129]. In other words, the behavior of markets is driven less by exogenous events and more by the endogenous dynamics of trading activity itself. The circular loop between price and trading – i.e., past price changes that feed on themselves – results in the erratic deviations from fundamentals, which are otherwise puzzling for standard economic theory but take a natural meaning when accepting the ubiquitous role of endogeneity.

In a series of studies, Sornette and collaborators have introduced measures of the degree of reflexivity or endogeneity in financial systems and built a theoretical framework that allows one to disentangle exogenous from endogenous sources of financial markets crashes [91,130–132]. In these studies, exogeneity refers to the external "forces" that influence the evolution of the system whereas endogeneity captures the self-reinforcing positive feedback processes within the system. Given the limited space, we can provide here only a brief overview of this research to give an intuition of the possible quantification of reflexivity. Combining statistical test of drawdowns distributions (runs of losses) and Log-Periodic Power Law Singular (LPPLS) detection techniques, Johansen and Sornette [133,134] showed that the extreme tails of the distribution belong to a different population than the body, analogue to the different physics that describe distributions in the study of turbulent hydrodynamic flows [135]. They then tested whether a LPPLS structure is present in the price trajectory, which precedes these "outliers," or "Dragon Kings" as Sornette [136] calls them (see also [137]). The emergence of log-periodic power law singular features is a qualifying signature of the endogenous dynamics, which might result in a market crash. As Sornette and co-workers has extensively documented, bubbles manifest themselves in super-exponential power-law accelerations in the price dynamics, which is decorated by log-periodic precursors. In their study, Sornette and Johansen [134] are able to identify two classes of market crashes: exogenously caused crashes that are not preceded by a LPPLS price trajectory and for which an exogenous shock can be identified, and crashes that are triggered endogenously by trading activity. The later are roughly twice as frequent as the former. Viewing crashes in terms of the endogenous dynamics of the market itself has important ramifications for our understanding of extreme event in financial systems. According to the view that most market crashes have an endogenous source – i.e., the increased cooperativity and self-organizing interactions between market participants – exogenous shocks only serve as triggering factors [135]. In other words, identifying proximal causal factors of a crash is often futile – as the extensive literature reflects, which presents diverse and often conflicting evidence about the origins of crashes (see [138]) – as the crash results from the maturation towards an intrinsically unstable phase.

In two more recent studies, our group has provided, for the first time, a quantification of reflexivity that allows us to precisely measure the levels of endogenity in a financial system [131,132]. For this, the so-called self-exciting Hawkes model is calibrated to financial market dynamics. This statistical model was initially developed to model earthquake clustering. As the Hawkes process formalism is able to describe the pattern of foreshocks and aftershocks, which result from the release of accumulated stress between tectonic plates, it is adaptable to financial markets where different regimes of volatility relaxation have been documented. For example, Sornette et al. [139] have shown that the relaxation time of a volatility burst is different after a strong exogenous shock compared with the relaxation of volatility after a peak with no identifiable exogenous source. They suggest that volatility can be understood in terms of response functions of financial agents, which derive from their behavior. By applying the Hawkes process analysis to E-mini S&P 500 futures, Filimonov and Sornette [131] are able to measure the degree of reflexivity as the proportion of price moves due to endogenous interactions to the total number of all price moves, which also include the impact of exogenous news. The self-exciting Hawkes branching process – in which each price changes may lead to an epidemic of other prices changes – allows one to identify different classes of volatility shocks along the separation between endogenous and exogenous dynamics. The Hawkes model has a key parameter, the "branching ratio" "n," – i.e., the fraction of endogenous events within the whole price-change population – that enables the direct measurement of the level of endogeneity. Interestingly, this measure reproduces robust behavioral feature of increased herding behavior at short time-scales in times of fear and panic. Filimonov and Sornette's [131] Hawkes process analysis of E-mini S&P 500 futures data from the period from 1998 to 2010 reveals a dramatic increase of endogeneity from 30% to 70–80% of trades triggered by past trades, an effect they attribute to the rise of high-frequency and algorithmic trading.

In a subsequent study on the endogeneity or reflexivity in commodity markets – using a Hawkes self-excited conditional Poisson model on time-series of past price-changes – Filimonov et al. [132] find that more than one out of two price changes is triggered by another price change, indicating a self-reinforcing reflexive mechanism underling the price time-series. Interestingly, Filimonov et al. [132] show that the level of endogeneity does not depend on the intensity of the information about exogenous events, which, as they document, has remained relatively stable over the analyzed period (second half of 20^{th} century to first decade of 21^{st} century). They further argue that increased reflexivity leads to a slower convergence of prices towards fundamental values, rendering the price formation process thereby less efficient. The study further shows that high levels of endogeneity or reflexivity result in a larger sensibility of the system to exogenous distortions. Endogenous feedback mechanisms in trading activity can amplify small initial shocks, which might, as it was the case with the May 6, 2010 flash crash, cascade into large crashes.

The studies on the endogenous versus exogenous sources of price volatility indicate that, far from being unquantifiable, reflexive financial phenomena can be disentangled and measured. This research seems to represent a first step towards the full quantification of reflexivity, from which novel insight about trader behavior, the evolution of bubbles, and the emergence of crashes will follow. While it is far from easy to get a quantitative grip on reflexivity [140][6], physics already possesses the concepts, tools, and methods needed to cope with reflexive phenomena. This has been shown in many other scientific fields involving complex systems. In principle, it can be concluded that the problem of reflexivity, as qualitatively described by Soros and many others, does not seem to inhibit the development of a new science of financial markets.

5 Conclusion

Historically, the collisions between physics and (financial) economics have resulted in the axiomatization of mainstream economics and gave rise to highly idealized models, which tend to be detached from financial reality. Although we do not deny that standard economic theory has generated deep insights into economic behavior, the neo-classical theoretical edifice with its fundamental pillars of rational expectations, efficiency, and equilibrium, nonetheless failed to deal with non-linear and out-of-equilibrium dynamics of complex financial systems, which are comprised of a myriad of heterogeneous agents interacting with each other. In particular, its failures became apparent to many during and after the great 2008 financial crisis [63,142,143].

Over the last two decades, however, econophysics has substantially advanced the scientific understanding of financial markets. The major contribution of econophysics was, as we have argued above, to view financial markets as complex systems. Describing markets in terms of emergence, scale-invariance, universality, and other properties of complexity, which we conceptually explicated above, allows one to better understand how the microscopic level – which is populated with mutually interacting agents that exhibit a diversity of behavioral traits, heterogeneous preferences and expectations – is linked with the macroscopic level of complexity, on which statistical regularities and patterns, i.e., the so-called stylized facts, can be identified. Econophysics has provided solid empirical foundations for the study of financial systems, which resulted in the falsification of many a priori assumptions of standard economic models. However, at the sociological level, the effectiveness of econophysics in the eyes of the economic profession has been limited due to econophysicists' disregard of standard economic theory and their misplaced aspiration to completely replace it with econophysics without due attention to previous achievements.

Although analogizations between physical and financial systems and extrapolations of concepts and methods of physics to (financial) economics can lead to oversimplified and idealized models of markets, physical concepts and techniques nonetheless provide a useful unifying framework to approach complexity. Contrary to the criticism that financial markets cannot be rendered intelligible by using the

[6] see [132,141] for recent technical developments and improvements.

scientific method of physics, we argued that, by integrating the insights from complex systems science, considerable progress has been made in the modeling of financial market dynamics. In fact, contrary to many claims to the contrary, we have argued that the problem of reflexivity does not demand a new scientific method. As we have documented above, research that applies the Hawkes process model, which was motivated by physics (or more precisely geophysics), to financial systems represents one of the first steps in the quantification of reflexive financial phenomena.

Nevertheless, it is difficult to argue that a new science of financial markets should solely have physics-based foundations. Understood in terms of the research program encoded in the term, econophysics, seems to be unnecessarily restrictive. At present, the most exciting progress seems to be unraveling at the boundaries between finance and the biological, cognitive, and behavioral sciences [144–146]. Recently, there have been many attempts that have started to explore the notion that financial markets are similar to ecologies, populated by species (traders, firms, etc.) that adapt and mutate. Farmer [147], for example, proposed a theory that views markets as ecologies in which – analogous to the evolution of a biological species – better-adapted strategies exploit old-strategies. Hommes [148] reviews the modeling of markets as evolutionary systems in which the survival of different trading strategies can be compared. Lo [149–151] proposed the "adaptive market hypothesis," which characterizes markets less in terms of efficiency, but rather in terms of competition and adaptation. Sornette [22] introduced the "Emerging Intelligence Market Hypothesis", according to which the continuous actions of investors, which are aggregated in the prices, produce a "market intelligence" more powerful than that of most of them. The "collective intelligence" of the market transforms most (but not all) strategies into losing strategies, just providing liquidity and transaction volume. Evolutionary models are able to explain most of the stylized facts documented in the econophysics literature (see [113]). These evolutionary or biological approaches provide another exiting source of inspiration for modeling financial reflexivity.

Concluding, econophysics should not be considered as isolated from other complexity-based approaches in science. What generates the statistical phenomena or patterns, which econophysics dissects with the tools of statistical physics, are fundamentally sociobiological systems. In other words, econophysics needs to reach beyond physics and integrate the concepts, methods, and tools from other disciplines. There can be a physics of financial markets, but the science we envision that helps to understand, diagnose, predict, and control financial markets [152,153] needs to integrate these fields into a complex systems science of financial markets.

References

1. E. Derman, *Models. Behaving. Badly: Why Confusing Illusion with Reality Can Lead to Disaster, on Wall Street and in Life* (Free Press, New York, 2011)
2. G. Soros, *The Alchemy of Finance: Reading the Mind of the Market* (John Wiley and Sons, Hoboken NJ, 1987)
3. E. Grunberg, F. Modigliani, J. Polit. Econ. **62**, 465 (1954)
4. H.A. Simon, Public. Open. Quart. **18**, 245 (1954)
5. R.C. Buck, Philos. Sci. **30**, 373 (1963)
6. M.K. Vetterling, Philos. Sci. **43**, 278 (1976)
7. E. Grunberg, Am. J. Econ. Sociol. **45**, 475 (1986)
8. G. Soros, J. Econ. Methodol. **20**, 309 (2013)
9. Y.C. Zhang, Physica A **269**, 30 (1999)
10. J.P. Bouchaud, Quant. Finance **1**, 105 (2001)
11. E. Derman, *My Life as a Quant: Reflections on Physics and Finance* (John Wiley and Sons, Hoboken NJ, 2004)

12. J.D. Farmer, T. Lux, J. Econ. Dyn. Control **32**, 1 (2008)
13. F. Jovanovic, C. Schinckus, Hist. Polit. Economy **45**, 443 (2013)
14. I. Newton, *The Principia: Mathematical Principles of Natural Philosophy: A New Translation* (University of California Press, Berkeley, 1999)
15. A. Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, edited by R.H. Campbell, A.S. Skinner, W.B. Todd (Oxford University Press, Oxford, 1976)
16. P.S. Laplace, in *Philosophical Essay on Probabilities: Translated from the Fifth French Edition of 1825 with Notes by the Translator*, edited by A.I. Dale (Springer, New York, 1995)
17. A. Quetelet, *A Treatise on Man and the Development of his Faculties*, edited by T. Smibert (Cambridge University Press, Cambridge 2013)
18. S. Galam, Y. Gefen, Y. Shapir, J. Math. Sociol. **9**, 1 (1982)
19. S. Galam, *Sociophysics: A Physicist's Modeling of Psycho-political Phenomena* (Springer, Berlin, 2012)
20. B. Ingrao, I. Israel, *The Invisible Hand: Economic Equilibrium in the History of Science* (MIT Press, Cambridge MA, 1990)
21. J.D. Farmer, J. Geanakoplos, Complexity **14**, 11 (2009)
22. D. Sornette, Rep. Prog. Phys. **77**, 062001 (2014)
23. V. Pareto, in *Oeuvres Complètes de Vilfredo Pareto*, edited by G. Busino (Librairie Droz, Geneva, 1981)
24. B. Mandelbrot, *The Fractal Geometry of Nature* (W.H. Freeman, San Francisco, 1982)
25. P. Bak, *How Nature Works: The Science of Self-organized Criticality* (Copernicus, New York, 1996)
26. D. Sornette, *Critical Phenomena in Natural Sciences, Chaos, Fractals, Self-organization and Disorder: Concepts and Tools*, 2nd edn. (Springer, Berlin, 2004)
27. R. Perline, Stat. Sci. **20**, 68 (2005)
28. A. Clauset, C.R. Shalizi, M.E.J. Newman, SIAM Rev. **51**, 661 (2009)
29. L. Bachelier, in *Louis Bachelier's Theory of Speculation: The Origins of Modern Finance*, edited by M. Davis, A. Etheridge (Princeton University Press, Princeton NJ, 2011)
30. A. Einstein, Ann. Phys. **322**, 549 (1905)
31. M.F.M. Osborne, Oper. Res. **7**, 145 (1959)
32. P.A. Samuelson, Indust. Manag. Rev. **7**, 41 (1965)
33. H. Markowitz, J. Finance **7**, 77 (1952)
34. W.F. Sharpe, J. Finance **19**, 425 (1964)
35. F. Black, M. Scholes, J. Polit. Economy **81**, 637 (1973)
36. R. Merton, Bell J. Econ. **4**, 141 (1973)
37. A.W. Lo, M.T. Mueller, J. Portfolio Manage. **8**, 13 (2010)
38. P.A. Samuelson, *Foundations of Economic Analysis* (Harvard University Press, Cambridge MA, 1947)
39. J.D. Farmer, J. Econ. Methodol. **20**, 377 (2013)
40. Y.C. Zhang, Europhys. News **29**, 51 (1998)
41. H.E. Stanley, L.A.N. Amaral, D. Canning, P. Gopikrishnan, Y. Lee, Y. Liu, Physica A **269**, 156 (1999)
42. H. Takayasu, H. Miura, T. Hirabayashi, K. Hamada, Physica A **184**, 127 (1992)
43. J.P. Bouchaud, D. Sornette, J. Phys. I France **4**, 863 (1994)
44. D. MacKenzie, *An Engine, Not a Camera: How Financial Models Shape Financial Markets* (MIT Press, Cambridge MA, 2006)
45. R.J. Shiller, Am. Econ. Rev. **71**, 421 (1981)
46. R.J. Shiller, *Market Volatility* (MIT Press, Cambridge MA, 1989)
47. S.F. LeRoy, R.D. Porter, Econometrica **49**, 555 (1981)
48. P. Krugman, *The Self-Organizing Economy*, 2nd edn. (Wiley-Blackwell, Oxford, 1996)
49. W.B. Arthur, Am. Econ. Rev. **84**, 406 (1994)
50. J. Voit, *The Statistical Mechanics of Financial Markets* (Springer, Berlin, 2005)

51. R.N. Mantegna, H.E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge, 2000)
52. R. Cont, Quant. Finance **1**, 223 (2001)
53. A. Chakraborti, I.M. Toke, M. Patriarca, F. Abergel, Quant. Finance **11**, 991 (2011)
54. J.P. Bouchaud, M. Potters, *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, 2nd edn. (Cambridge University Press, Cambridge, 2003)
55. B. Mandelbrot, Quart. J. Econ. **76**, 57 (1962)
56. B. Mandelbrot, J. Bus. **36**, 394 (1963)
57. B. Mandelbrot, J. Bus. **39**, 242 (1966)
58. E.F. Fama, J. Bus. **36**, 420 (1963)
59. E.F. Fama, Manag. Sci. **11**, 404 (1965)
60. P.A. Samuelson, J. Finan. Quant. Anal. **2**, 107 (1967)
61. J.P. Bouchaud, D. Sornette, C. Walter, J.P. Aguilar, Int. J. Theor. Appl. Finance **1**, 25 (1998)
62. P.H. Cootner, *The Random Character of Stock Market Prices* (MIT Press, Cambridge MA, 1964)
63. D. Colander, M. Goldberg, A. Haas, K. Juselius, A. Kirman, T. Lux, B. Sloth, Crit. Rev. **21**, 249 (2009)
64. J.L. McCauley, *Dynamics of Markets: Econophysics and Finance* (Cambridge University Press, Cambridge, 2004)
65. S. Keen, Physica A **324**, 108 (2003)
66. M. Gallegati, S. Keen, T. Lux, P. Ormerod, Physica A **370**, 1 (2006)
67. P. Ball, Nature **441**, 686 (2006)
68. J.B. Rosser, J. Econ. Perspect. **13**, 169 (1999)
69. T.C. Schelling, Am. Econ. Rev. **59**, 488 (1969)
70. T.C. Schelling, J. Math. Sociol. **1**, 143 (1971)
71. T.C. Schelling, *Micromotives and Macrobehavior* (W.W. Norton, New York, 1978)
72. R. Frigg, J. Reiss, Synthese **169**, 593 (2009)
73. J.M. Epstein, *Generative Social Science: Studies in Agent-Based Computational Modeling* (Princeton University Press, Princeton NJ, 2006)
74. Z. Burda, J. Jerzy, M.A. Nowak, Acta Phys. Pol. B **34**, 87 (2003)
75. K. Arrow, P. Anderson, D. Pines (eds.), *The Economy as an Evolving Complex System* (Addison Wesley, Reading MA, 1988)
76. W.B. Arthur, S. Durlauf, D. Lane (eds.), *The Economy as an Evolving Complex System II* (Addison Wesley, Reading MA 1997)
77. L. Blume, S. Durlauf (eds.), *The Economy as an Evolving Complex System III* (Oxford University Press, New York, 2006)
78. P.W. Anderson, Science **177**, 393 (1972)
79. J.M. Coates, *The Hour Between Dog and Wolf: Risk Taking, Gut Feelings and the Biology of Boom and Bust* (Penguin Press, New York, 2012)
80. D. Sornette, *Why Stock Markets Crash: Critical Events in Complex Financial Systems* (Princeton University Press, Princeton NJ, 2003)
81. T. Kaizoji, D. Sornette, in *Encyclopedia of Quantitative Finance*, edited by R. Cont (John Wiley and Sons, Hoboken NJ, 2010) (long version at `http://arXiv.org/abs/0812.2449`)
82. S. Durlauf, Econ. J. **115**, F225 (2005)
83. W.B. Arthur, in *Complexity and the Economy*, edited by W.B. Arthur (Oxford University Press, New York, 2014)
84. D.A. Hsieh, J. Bus. **62**, 339 (1989)
85. D.A. Hsieh, Financ. Anal. J. **51**, 55 (1995)
86. D. Sornette, P. Cauwels, Rev. Behav. Econ. **2**, 279 (2015)
87. A. Hüsler, D. Sornette, C.H. Hommes, J. Econ. Behav. Organ. **92**, 304 (2013)
88. T. Kaizoji, M. Leiss, A. Saichev, D. Sornette, J. Econ. Behav. Org. **112**, 289 (2015)
89. M. Leiss, H.H. Nax, D. Sornette, J. Econ. Dyn. Control **55**, 1 (2015)

90. D. Sornette, Physica A **284**, 355 (2000)
91. A. Johansen, D. Sornette, O. Ledoit, J. Risk **1**, 5 (1999)
92. A. Johansen, O. Ledoit, D. Sornette, Int. J. Theor. Appl. Finance **3**, 219 (2000)
93. Z.Q. Jiang, W.X. Zhou, D. Sornette, R. Woodard, K. Bastiaensen, P. Cauwels, J. Econ. Behav. Organ. **74**, 149 (2010)
94. D. Sornette, R. Woodard, W. Yan, W.X. Zhou, Physica A **392**, 4417 (2013)
95. D. Sornette, P. Cauwels, J. Risk Manag. Finan. Institutions **8**, 83 (2015)
96. D. Sornette, Phys. World **15**, 50 (2002)
97. D.T. Campbell, in *Studies in the Philosophy of Biology*, edited by F.J. Ayala, T. Dobzhansky (University of California Press, Berkley, 1974)
98. G.F.R. Ellis, T. Roy. Soc. S. Afr. **63**, 69 (2008)
99. U. Frisch, *Turbulence: The Legacy of A. N. Kolmogorov* (Cambridge University Press, Cambridge, 1996)
100. N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (Addison-Wesley, Boston, 1992)
101. D. Delli Gatti, E. Gaffeo, M. Gallegati, G. Giulioni, A. Palestrini, *Emergent Macroeconomics: An Agent-Based Approach to Business Fluctuations* (Springer, Milan, 2008)
102. R. Cont, in *Fractals in Engineering*, edited by J. Lévy-Véhel, E. Lutton (Springer, London, 2005)
103. M. Gallegati, in *Econophysics of Markets and Business Networks*, edited by B.K. Chakrabarti, T. Kaizoji, Y. Wang, Y. Fujiwara, S. Sinha, J. Mimkes (Springer, Milan, 2007)
104. W. Hildebrand, A. Kirman, *Equilibrium Analysis: Variations on the Themes by Edgeworth and Walras* (North-Holland, Amsterdam, 1988)
105. D. Sornette, P. Cauwels, Risks **2**, 103 (2014)
106. M. Friedman, *Essays in Positive Economics* (University of Chicago Press, Chicago, 1953)
107. D.M. Cutler, J.M. Poterba, L.H. Summers, J. Portfolio Manage. **15**, 4 (1989)
108. D.M. Cutler, J.M. Poterba, L.H. Summers, Rev. Econ. Stud. **58**, 529 (1991)
109. J.A. Schumpeter, *Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process* (McGraw-Hill, New York, 1939)
110. P. Mirowski, *More Heat than Light: Economics as Social Physics, Physics as Nature's Economics* (Cambridge University Press, Cambridge 1991)
111. C. Schinckus, Quant. Finance **12**, 1189 (2012)
112. C. Schinckus, Contemp. Phys. **54**, 17 (2013)
113. A. Chakraborti, I.M.T. Patriarca, F. Abergel, Quant. Finance **11**, 991 (2011)
114. P. Bak, M. Paczuski, M. Shubik, Physica **A 246**, 430 (1997)
115. S. Mike, J.D. Farmer, J. Econ. Dynam. Control **32**, 1 (2008)
116. Y. Yura, H. Takayasu, D. Sornette, M. Takayasu, Phys. Rev. Lett. **112**, 098703 (2014)
117. Y. Yura, H. Takayasu, D. Sornette, M. Takayasu, Phys. Rev. E **92**, 042811 (2015)
118. J.D. Farmer, P. Patelli, I.I. Zovko, Proc. Natl. Acad. Sci. USA **102**, 2254 (2005)
119. B. LeBaron, in *Handbook of Computational Economics Volume II: Agent-Based Computational Economics*, edited by L. Tesfatsion, K.L. Judd (North-Holland, Amsterdam, 2006)
120. T. Lux, M. Marchesi, Nature **397**, 498 (1999)
121. T. Lux, M. Marchesi, Int. J. Theor. Appl. Finance **3**, 675 (2000)
122. T. Lux, Working Paper No. 1425, University of Kiel, 2008
123. W.X. Zhou, D. Sornette, Eur. Phys. J. B **55**, 175 (2007)
124. G. Harras, D. Sornette, J. Econ. Behav. Organ. **80**, 137 (2011)
125. H.A. Simon, in *Proceedings of the American Philosophical Society* **106**, 467 (1962)
126. D. Sornette, in *Extreme Events in Nature and Society*, edited by S. Albeverio, V. Jentsch, H. Kantz (Springer, Berlin, 2005)
127. A. Joulin, A. Lefevre, D. Grunberg, J.P. Bouchaud, Wilmott Magazine **46** (Sept./Oct. 2008)

128. J.P. Bouchaud, in *Lessons from the 2008 Crisis*, edited by A. Berd (Risk Books, Incisive Media, London, 2011)
129. J.P. Bouchaud, Procedia Comput. Sci. **7**, 22 (2011)
130. A. Johansen, D. Sornette, in *Contemporary Issues in International Finance*, edited by F. Columbus (Nova Science, New York, 2004)
131. V. Filimonov, D. Sornette, Phys. Rev. E **85**, 056108 (2012)
132. V. Filimonov, D. Bicchetti, N. Maystre, D. Sornette, J. Int. Money Finance **42**, 174 (2014)
133. A. Johansen, D. Sornette, J. Risk **4**, 69 (2001)
134. A. Johansen, D. Sornette, Cah. Econ. Bruxelles **53**, 201 (2010)
135. D. Sornette, Phys. Rep. **378**, 1 (2003)
136. D. Sornette, Int. J. Terraspace Sci. Eng. **2**, 1 (2009)
137. D. Sornette, G. Ouillon, Eur. Phys. J. Special Topics **205**, 1 (2012)
138. A.W. Lo, J. Econ. Lit. **50**, 1 (2012)
139. D. Sornette, Y. Malevergne, J.F. Muzy, Risk **16**, 67 (2003)
140. V. Filimonov, D. Sornette, Quant. Finance **15**, 1293 (2015)
141. S. Wheatley, V. Filimonov, D. Sornette, Comput. Stat. Data Analysis **94**, 120 (2016)
142. P. Krugman, Eastern Econ. J. **37**, 307 (2011)
143. D.J. Bezemer, J. Econ. Issues **45**, 1 (2011)
144. C.F. Camerer, G. Loewenstein, M. Rabin (eds.), *Advances in Behavioral Economics* (Princeton University Press, Princeton NJ, 2003)
145. R.J. Shiller, J. Econ. Perspect. **17**, 83 (2003)
146. R.H. Thaler (ed.), *Advances in Behavioral Finance: Volume II* (Princeton University Press, Princeton NJ, 2005)
147. J.D. Farmer, Indust. Corp. Change **11**, 895 (2002)
148. C.H. Hommes, Quant. Finance **1**, 149 (2001)
149. A.W. Lo, J. Portfolio Manag. **30**, 15 (2004)
150. A.W. Lo, J. Invest. Consulting **7**, 21 (2005)
151. A.W. Lo, Financ. Anal. J. **68**, 18 (2012)
152. S. Cincotti, D. Sornette, P. Treleaven, S. Battiston, G. Caldarelli, C. Hommes, A. Kirman, Eur. J. Phys. Special Topics **214**, 361 (2012)
153. H.L.D. de S. Cavalcante, M. Oriá, D. Sornette, E. Ott, D.J. Gauthier, Phys. Rev. Lett. **111**, 198701 (2013)

3. Bubbles and Technological Innovation: Bitcoin Case Study

Full Reference:

Huber, T. and Sornette, D. (2020), Boom, Bust, and Bitcoin: Bitcoin-Bubbles As Innovation-Accelerators." Forthcoming *Journal of Economic Issues.*

Abstract

Bitcoin represents one of the most interesting technological breakthroughs and socio-economic experiments of the last decades. In this paper, we examine the role of speculative bubbles in the process of Bitcoin's technological adoption by analyzing its social dynamics. We trace Bitcoin's genesis and dissect the nature of its techno-economic innovation. In particular, we present an analysis of the techno-economic feedback loops that drive Bitcoin's price and network effects. Based on our analysis of Bitcoin, we test and further refine the Social Bubble Hypothesis, which holds that bubbles constitute an essential component in the process of technological innovation. We argue that a hierarchy of repeating and exponentially increasing series of bubbles and hype cycles, which has occurred over the past decade since its inception, has bootstrapped Bitcoin into existence.

s:fi RESEARCH

Swiss Finance Institute
Research Paper Series
N°20-41

Boom, Bust, and Bitcoin: Bitcoin-Bubbles As Innovation Accelerators

Tobias Huber
ETH Zurich

Didier Sornette
ETH Zurich, Southern University of Science and Technology (SUSTech) Shenzhen, Tokyo Institute of Technology, and Swiss Finance Institute

Boom, Bust, and Bitcoin:
Bitcoin-Bubbles As Innovation Accelerators

Tobias A. Huber[1] and Didier Sornette[1-3]

[1]ETH Zurich, Department of Management, Technology and Economics, Switzerland
[2]Tokyo Tech World Research Hub Initiative, Institute of Innovative Research, Tokyo Institute of Technology
[3]Institute of Risk Analysis, Prediction and Management (Risks-X), Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology (SUSTech), Shenzhen, 518055, China

May 12, 2020

Abstract: Bitcoin represents one of the most interesting technological breakthroughs and socio-economic experiments of the last decades. In this paper, we examine the role of speculative bubbles in the process of Bitcoin's technological adoption by analyzing its social dynamics. We trace Bitcoin's genesis and dissect the nature of its techno-economic innovation. In particular, we present an analysis of the techno-economic feedback loops that drive Bitcoin's price and network effects. Based on our analysis of Bitcoin, we test and further refine the Social Bubble Hypothesis, which holds that bubbles constitute an essential component in the process of technological innovation. We argue that a hierarchy of repeating and exponentially increasing series of bubbles and hype cycles, which has occurred over the past decade since its inception, has bootstrapped Bitcoin into existence.

Keywords:
Bitcoin, Money, Cryptocurrencies, Financial Bubbles, Technological Innovation, Economic Growth, Reflexivity

JEL: E40; G01; O40

1. Introduction

The emergence of Bitcoin represents one of the most interesting technological breakthroughs and socio-economic experiments of the last decades. Not only is Bitcoin[1] a multi-faceted object, which synthesizes techno-scientific insights from various fields, it also triggered one of the largest speculative bubbles in human history. However, while the underlying cryptography, game-theory, or monetary economics, for example, have attracted scientific interest and generated an expanding academic literature,[2] a comprehensive analysis of Bitcoin as a socio-economic innovation—which situates the phenomenon in the context of the history and economics of technological innovation—has so far been lacking. In this paper, we will focus on bitcoin bubbles, in which the social dynamics driving the development and adoption of the technology crystallize themselves.

By combining various technical components, such as peer-to-peer network technology, asymmetric public-key cryptography, and a new proof-of-work algorithm, Bitcoin represents a novel, decentralized digital form of money, which substantially reduces the need for trusted third parties. As envisioned by its creator and many of his followers, it represents a radical technological innovation with potentially far-reaching socio-economic consequences. Now, irrespective of whether this vision will get realized or fail, Bitcoin offers a historically singular opportunity to advance our understanding of the nature of technological innovation and its socio-economic effects. As it is an extraordinary natural experiment—which is highly instructive but, at this scale, extremely rare in economics—the study of Bitcoin can illuminate, as we will demonstrate here, the nature of technological innovation and the social dynamics that catalyze the diffusion of emerging technologies.

Invented by a pseudonymous programmer and introduced online on an obscure mailing list, Bitcoin has since its inception experienced hyperbolic growth. Whereas the network has grown from one to an estimated 9400 nodes, the market capitalization of the protocol's cryptocurrency hyperbolically exploded from zero to almost 300 Billion USD at the peak of the so-called "crypto mania" in early 2018. As we will argue in this paper, a critical component in bootstrapping and scaling the Bitcoin network was the bubble-sequence that has driven its adoption. By accelerating the cryptocurrency's price, which, in turn, has catalyzed its speculative adoption, a series of spectacular bubbles has bootstrapped the Bitcoin network into existence. In order now to elucidate how financial speculation and excessive hype can function as necessary components in the process of developing, adopting, and diffusing novel technologies, we will dissect in the following the interwoven technological, economic, and social feedback loops that fuel Bitcoin. As we will show, Bitcoin is equally a technical as well as a social phenomenon—it represents a revolutionary technological breakthrough that will potentially have radical future socio-economic consequences. Thus, it provides a unique occasion to

[1] In this paper, we follow the convention to use uppercase "B" to refer to the Bitcoin network and lowercase "b" to the protocol-native cryptocurrency.

[2] For a sample of the academic literature, see Narayanan et al., 2016, Narayanan and Clark, 2017, Usman 2017, Wheatley et al., 2019, and Gerlach et al., 2019. However, most of the relevant analyses, on which we will rely, have occurred outside traditional academic journals on blogs or social networks, such as Medium, Reddit, Twitter, or Telegram.

test and extend the Social Bubble Hypothesis, which has been developed in our research group (see Sornette, 2008; Gisler and Sornette, 2009; 2010; Gisler et al., 2011; Huber, 2017). In essence, the hypothesis, which we have refined over a series of case studies, holds that bubbles are necessary elements in the social, economic, and political processes that result in large-scale and high-impact innovations. However, while we will provide an overview of the technical properties and design of the Bitcoin network, our analysis will focus here on the socio-economic nature of Bitcoin. Consequently, the aim of the paper is threefold: i) to illuminate the technological, social, and economic dimensions of Bitcoin and their interactions; ii) to improve the hypothesis that bubbles can be phenomena with net positive benefits, as they can incubate technological and societal change, and; (iii) to extract valuable and generalizable insights from the history of Bitcoin that might help us to advance our understanding of the generic dynamics and structure of future technological revolutions.

While financial bubbles and market crashes have attracted abundant attention in quantitative finance and financial economics, research in these fields has mainly focused on bubbles as negative phenomena, which instantiate a form of economic inefficiency or market failure (see Allen and Gorton, 1993; Santos and Woodford, 1997; Abreu and Brunnermeier, 2003; Scheinkman and Xiong, 2003; Garber, 2000). Countering this view of bubbles as economically destructive or unproductive economic phenomena, new research has, over the last two decades, emerged outside the academic mainstream in economics that conceptualizes bubbles as important components in the process of techno-social innovation (Perez 2002; Sornette, 2008; Janeway, 2012). By analyzing Bitcoin through the conceptual prism of the Social Bubble Hypothesis, we aim to further develop our unified theory of financial bubbles and crashes (see Sornette, 2003), which provides a scientific foundation for diagnosing and predicting financial bubbles and crashes and incorporates a generic explanation for their role in technological innovation. By examining the evolution of the protocol and its cryptocurrency, we will show in more detail below that the integration of previously existing technical components and ideas into the design of the protocol—which allowed to bridge disparate and unrelated fields, methods, and concepts—enabled Bitcoin's technological breakthrough. In other words, Bitcoin provides, in our view, a valuable blueprint that can help us better understand and anticipate future technological innovations. While our analysis cannot exhaustively capture the multidisciplinary nature of Bitcoin, we nevertheless aim to identify in this paper the essential technological, social, and economic dynamics that accelerate the development and dissemination of Bitcoin.

The paper is structured as follows. In the first part, we will trace Bitcoin's historical evolution and give a synoptic view of the protocol and the technical properties of the cryptocurrency. In this section, we isolate the process of technological innovation that has resulted in Bitcoin's breakthrough. We will argue that Bitcoin instantiates a form of "combinatorial evolution," which captures the process from which novel technologies arise from the combination of preceding technological elements (see Arthur, 2009). As its historical genesis shows, Bitcoin represents an instance of "radical" or "vertical" technological innovation—as opposed to "incremental" or "horizontal" progress—precisely because its "novelty" emerged from a assemblage of existing technological components and economic incentives (see Thiel and Masters, 2012). In the next section, we will examine the social dimension of Bitcoin. By dissecting the incentive system that is embedded in the protocol, we will show that Bitcoin solves a large-scale social coordination problem by automating and formalizing social consensus between network participants and economic agents. We then analyze the cultural forces that are shaping Bitcoin's development and adoption. In particular, we identify the guiding visions driving Bitcoin's development and evangelism and the subcultures that have formed around them. Our analysis of the belief-systems that have emerged around the protocol and cryptocurrency reveals the quasi-religious dimension of Bitcoin, which manifests itself, for example, in the beliefs of some of the most committed supporters and their exegesis of Nakamoto's code and writings. As they share a structural similarity with self-fulfilling prophecies (see Merton, 1948), the beliefs that incentivize technical development and financial speculation in Bitcoin are, as we will argue, a critical factor in understanding the technology's rapid diffusion, increase in network-security, and appreciation in value. In the third section, we will further explore the self-validating nature of the Bitcoin phenomenon by mobilizing the Social Bubble Hypothesis. Based on an examination of the hype-cycles that punctuate the history of Bitcoin, we will then develop and substantiate the argument that bitcoin bubbles were necessary to bootstrap and scale the protocol and cryptocurrency. We propose that these exuberant phases need to be conceptualized as instances of *speculative technology adoption* (see also Casey, 2016). In the last section, we conclude the paper by discussing how Bitcoin can enlighten our understanding of future technological revolutions more generally.

2. Bitcoin: From Zero to One

On January 8, 2009, the pseudonymous programmer Satoshi Nakamoto released, on an obscure cryptography mailing list, Bitcoin—a software-protocol that allows for the decentralized transmission and storage of value. In his 2008 white paper, which provides the conceptual blueprint of the Bitcoin network, Nakamoto characterizes it as "a new electronic cash system that's fully peer-to-peer, with no trusted third party" (Nakamoto, 2008). However, the path that led to decentralized and peer-to-peer digital money is littered with failed attempts. Historically, two main currents can be identified that have coalesced into Bitcoin. Tracing back the influence on Nakamoto's protocol, these currents derive from two intertwined historical developments: cryptographic advancements in computer science and the ideologically-motivated development of cryptographically-secured, non-sovereign virtual currencies.[3] In this section, we will provide a brief outline of the intellectual history of the technical and crypto-anarchic ideas preceding Bitcoin's invention. An understanding of its pre-history will help us recognize the significance of Bitcoin as a technological as well as socio-economic breakthrough.

2.1. Bitcoin: A Selective History

Nakamoto's invention is preceded by many failed attempts to create virtual currencies, such as DigiCash, Hashcash, or Bitgold. One of the earliest and most prominent proposals is David Chaum's DigiCash. In 1989, Chaum founded DigiCash, which—by applying public key cryptography to the specific problem of

[3] This sections draws on the exceedingly thorough review of Bitcoin's academic pre-history by Narayanan and Clark (2017).

digital monetary transactions— attempted to create cryptographically-secured digital cash, which emulated the properties of physical money. As early as 1983, he published the paper "Blind Signatures for Untraceable Cash" (see Chaum, 1983), which proposed so-called "blind signatures" that enabled privacy in transactions and avoided the "double-spending"-problem, which plagued many early attempts of creating digital cash. DigiCash developed a currency called ecash, which was an untraceable system of digital cash. While some banks implemented ecash, and Microsoft even proposed to integrate it into Windows, DigiCash ultimately declared bankruptcy in 1998 because of lack of merchant-adoption and support of user-to-user transactions. A wave of digital payment startups—with generic names such as CyberGold, CyberCash, or E-Gold—followed DigiCash's invention of ecash in the mid-1990s, in attempts to develop web-based money. However, except for Paypal—which pivoted away from their initial idea of enabling cryptographic payments through handheld Palm Pilot devices—these online-payment startups failed. Parallel to DigiCash and other attempts to patent and commercialize digital currencies and online-payment systems, a group of cryptographers, who interacted on what was called the "cypherpunk" mailing list, started to develop open-source alternatives (see Narayanan and Clark, 2017). While some projects, such as e-gold, proposed to peg the value of digital cash to a fiat currency or commodity, others started to experiment with free-floating digital currencies. By simulating the properties of gold, for example, some of these proposals attempted to digitally re-engineer gold's scarcity as a source of value for the native virtual currencies of these payment networks. In Bitcoin, digital scarcity has been achieved by designing a payment architecture in which the creation of money requires solving computationally expensive problems. Ideas for such systems, which Bitcoin later implemented with its proof-of-work algorithm, date back to a proposal by cryptographers Cynthia Dwork and Moni Naor, which was published in the early 1990s (see Dwork and Naor, 1992). The term *proof of work* was coined in a paper by Jakobsson and Juels in 1999 (see Jakobsson and Juels, 1999). In their paper, Dwork and Naor proposed a system in which the solution of computational problems (or "puzzles") was used to reduce email-spam. A similar idea later was expressed in Adam Back's Hashcash proposal, which he published in 1997.

In proof-of-work systems, which Hashcash and similar proposals have pioneered, the transaction validation and associated digital currency issuance—the work that needs to be proven—is performed by CPUs that invest computational recourses into a mathematical puzzle-solving exercise.[4] Although the name Hashcash already implicitly contains the idea of monetizing proof-of-work certification, the Hashcash stamps themselves, which constitute the computational proofs-of-work, were not designed to acquire monetary value. A member of the techno-libertarian cypherpunk community, Back envisioned Hashcash's proof of work-system as digital cash, and thus as an alternative to Chaum's Digicash. However, it was not possible to

[4] While it is tempting to construe proof-of-work as an algorithmic reformulation of the labor theory of value—which postulates that value is determined by labor or the cost of production—"work" in Bitcoin's system is derived not from political economy but from computer science. The work to be proven—that is, transaction validations and cryptocurrency issuance performed by CPUs—is probabilistic and not deterministic in nature. In other words, no amount of computational effort guarantees a reward. Rather, the successful solution of a cryptographic puzzle is a low-probability outcome, which miners try to achieve in repeating trial-and-error-processes (see Land, forthcoming). More generally, an adequate economic framework for understanding the process of bitcoin's monetization is not the Marxist labor theory of value—elements of which can be already identified in Aristotle, Adam Smith, or David Ricardo—but the Austrian monetary economics developed by Carl Menger, Ludwig Van Mises, or F.A. Hayek. For an application of Austrian economics to Bitcoin, see Ammous (2018).

exchange Hashcash stamps across a peer-to-peer network. Developed after Back's Hashcash proposal, more developed proposals, which conceptualize computational puzzle solutions as digital cash, have been advanced with b-money and Bitgold (see Dai, 1998; Szabo, 2008). In both proposals, the process of solving computational puzzles is directly used for the production of digital currency. In Bitgold and b-money, which both use time-stamping to validate transactions, the computational solutions themselves instantiate monetary units. However, b-money and Bitgold, which were informally proposed on a mailing list and in a series of blog posts respectively, did not advance beyond the conceptual stage of development—they were both not implemented and lacked any code specifications.

While Nakomoto stated in 2010 on the Bitcointalk.org forum that "Bitcoin is an implementation of Wei Dai's b-money proposal on Cypherpunks in 1998 and Nick Szabo's Bitgold proposal," Bitcoin represents a technological novelty as it goes much beyond just implementing a set of pre-existing cryptographic ideas. Instead, the design of Bitcoin's architecture specifically solved deep technical and conceptual issues that earlier proposals for digital currency systems had not fully fleshed out or simply failed to address. More specifically, Hashcash, Bitgold, or b-money were all undermined by two core problems that Nakamoto's design solved: the self-monetization of the protocol-native cryptocurrency and the decentralization of network governance.

Whereas the Hashcash-system, for example, critically lacked any control of inflation, Bitcoin incorporates an automatic mechanism to periodically adjust the difficulty of the computational puzzles that regulate the issuance of new cryptocurrency. Thus, Bitcoin is capable of responding to declining hardware costs for a fixed amount of computing power, which—by substantially lowering the difficulty of producing a cryptocurrency—would result in its devaluation. By adopting an upgraded version of the Hashcash algorithm for the Bitcoin mining process, the mining difficulty adjustment—which governs Bitcoin's proof-of-work system—solved the inflation control problem, which plagued many previous digital cash proposals. In other words, as it automatically adjusts the difficulty to stabilize the rate of supply, Nakamoto was able to design a decentralized form of digital money, which removes the need for any central authority to control the inflation rate or secure the network. The monetary policy that is embedded in the protocol—which caps its supply at 21 million bitcoins—and the difficulty adjustment, which regulates the flows of energy being expended to secure the network, are, as we will show in more detail below, the source of Bitcoin's technological innovation. Furthermore, both Bitgold and b-money, for example, did not specify a consensus-mechanisms to resolve disagreement among nodes or servers about the ledger that stores all transactions in the network. Settling disagreements would have then required trusted time-stamping services for currency-creation and validation, and centralized entities controlling entry into the network to secure it from attackers attempting to alter the ledger's history or double-spend the virtual monetary units.[5] In Bitcoin, this problem is solved by the so-called "mining"-process, which was intentionally designed to be resource-intensive and computationally difficult so as to ensure that the number of mined blocks—which record transaction data—

[5] In a 2009 post on the P2P Foundation message board, Nakamoto states: "A lot of people automatically dismiss e-currency as a lost cause because of all the companies that failed since the 1990's. I hope it's obvious it was only the centrally controlled nature of those systems that doomed them. I think this is the first time we're trying a decentralized, non-trust-based system" (Nakamoto, 2009).

remains steady. Instead of relying on trusted servers that time-stamp transactions in a ledger—as it was proposed, for example, in a series of academic papers by Haber and Stornetta in the 1990s (see Haber and Stornetta, 1991)—Bitcoin transactions are collected by a network of untrusted "miners," which are compensated with new bitcoins and transaction fees to permanently record transaction-data into irreversible "blocks." Ordered in a linear sequence, these block give rise to what Nakamoto called "time-chain," which, later, became popularized as "blockchain" (see Nakamoto, 2009).

As this highly selective and brief history demonstrates, the design of Bitcoin synthesizes a set of existing core technical elements. Public key cryptography, Merkle Trees, cryptographic signatures and hash functions, proof-of-work, and other insights derived from the engineering of resilient peer-to-peer networks in computer science have provided the building material for the architecture of the Bitcoin network. While we are not going to delve into the intricate details of Bitcoin's technical properties[6], it is, for the purpose of this paper, sufficient to understand the network's key-components, which were incubated in its academic and cypherpunk predecessors, in order to recognize how their novel combination gave rise to Bitcoin's radical technological innovation. In order to better appreciate the breakthrough that Nakamoto's design represents, we need to zero in on the reflexive feedback loops that drive Bitcoin's security, value, and network effects. Consequently, we briefly dissect in the next section the structure of the technological as well as socio-economic incentives that are built into the protocol.

2.2 The Techno-Economic Reflexivity of Bitcoin

As the previous section indicates, Bitcoin represents not only a material but, in a fundamental sense, also a social technology. While the material technology that underlies the Bitcoin network consists of its codebase, the physical mining rigs, or nodes that run the Bitcoin Core software, developers, speculators, or miners, for instance, constitute the social layer of the Bitcoin architecture. Bitcoin as a social technology coordinates the behavior of this heterogeneous group of network participants that is needed for the governance of the protocol. In other words, Bitcoin's protocol governance denotes a fundamentally social process that decides upon, implements, and enforces a set of transaction and block-verification rules, which network participants can adopt. By adopting the same set of validation rules, network participants form an inter-subjective consensus about what constitutes "Bitcoin" (see Rochard, 2018). Dissenting network participants can only deviate from this inter-subjective definition of Bitcoin by "hard-forking" the protocol, that is, upgrading a copied version of the software to a new set of transaction- and block-verification rules or a different blockchain history.

At the core of the Bitcoin system-architecture, we can identify two reflexive components of a self-validating positive *techno-socio-economic feedback loop*, which incentivizes Bitcoin's development, valuation, and adoption (see Figure 1)[7]:

---

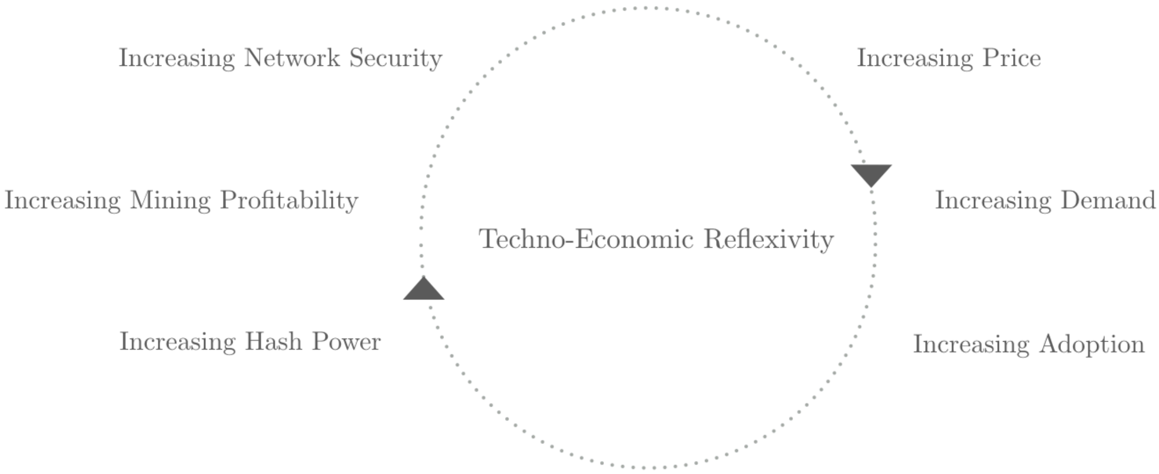
[6] For technical treatments of Bitcoin, see Antonopoulos (2014); Song (2019).

[7] As the bust of the bitcoin bubble in 2018 has shown, the positive feedback loop underlying bitcoin's rise can revert into a dynamic that results in accelerating price and devaluations accompanied respectively by increases and decreases in hash power, which, in turn, reduce network activity and security.

Technological Reflexivity: As mentioned above, specialized "miners" secure, maintain, and issue new bitcoin by competing to solve computationally intensive cryptographic puzzles. In Bitcoin's proof-of-work system, the probability of success of miners—which can be organized as companies or as mining pools—is determined by the fraction of mining power they control. If a miner successfully solves a puzzle, it gets to contribute the next "block" of transactions to the blockchain, in which blocks of transactions are linked together based on time-stamping. The mining entities are incentivized to secure and maintain the network with newly issued bitcoins. As block rewards for invalid transactions or blocks will be invalidated and rejected by the majority of miners, miners' incentives to comply with the protocol and its rules are aligned. Bitcoin's design avoids the double-spending problem, which has plagued earlier digital currency proposals, as the puzzle solutions themselves are decoupled from economic value. The amount of work required to produce a block and the amount of bitcoins issued are not fixed parameters. Rather, the block reward—which is at the time of writing (March 2020) 12.5 bitcoins/block—is programmed to halve every four years, or, approximately every 210,000 blocks (the next halving is expected to occur in May, 2020). Beyond the block reward by which new bitcoins are generated, miners are incentivized to maintain and secure the network by an additional reward scheme embedded into Bitcoin's design: senders of bitcoin payments pay miners a fee for their service of including the transaction into a block. This combined reward system that Nakamoto hard-coded into the protocol fuels, on the technological level, the reflexive feedback loop between the networks' security and value. Growth in bitcoin's value results in increased hash power allocated to the network, which, in turn, enhances its security and attracts new miners and the development and deployment of specialized mining hardware.

Socio-Economic Reflexivity: The reflexive incentive loop, which is built into the protocol layer and governs bitcoin's technological development, gives rise to a feedback loop on the socio-economic level. As the network's security and the native cryptocurrency's value increase, speculators, investors, and entrepreneurs are incentivized to explore and exploit the new economic space opened up by Bitcoin. Bitcoin's growth over the past decade has been driven by this self-reinforcing dynamic between discovery and speculation, which can be identified as a generic pattern in the major technological innovations of the past 250 years (see Perez, 2002; Janeway, 2013). Massive price-increases, which have been fueled by speculation, have triggered successive processes of experimentation that resulted in the gradual build-out of Bitcoin's economic infrastructure, such as the development of the second layer payment-system Lightning Network, Bitcoin-based startups that provide services such as brokerage, exchanges, wallets, public key storage, or novel alternative Bitcoin transaction systems, such as mesh or satellite networks, which are not relying on traditional Internet Service Providers.

The interlocking positive spirals of technological development and economic expectations, in turn, induce a self-validating, imitation-driven social feedback loop that accelerates Bitcoin's network effect. As past bubbles have demonstrated, exploding prices generate interest and attention, as has occurred during Bitcoin's hype cycles. Bitcoin's extraordinary returns, and the resulting media coverage and contagious virality on social media, elicit the "fear-of-missing-out" future gains inferred from extrapolations of the great gains that others have accrued in the recent past. This leads to new waves of buyers that accelerate the price growth even more. This increase in speculation and adoption, in turns, stimulates the founding and financing of new Bitcoin-based, or other cryptocurrency-related, startups that attract more speculative adopters and incentivize more Bitcoin-related research and development.



In essence, these nested reflexive feedback loops, which constitute the incentive-system underlying Bitcoin's design, reveal the technological as well as socio-economic novelty of Nakamoto's invention. Incentivized by the scarcity encoded in the protocol, increasing demand results in more investments into specialized mining hardware and hash power allocation to the network, which, in turns, attracts more speculators, miners, and entrepreneurs that accelerate the self-validating reflexive feedback loop of security, value, and network effects.[8] In other words, Bitcoin represents a circular or closed system of socio-techno-economic incentives. As a forum post on the P2P Foundation website, dated February 18, 2009, indicates, Nakamoto himself was fully aware of the reflexive feedback loop that the design of the protocol incubates: "As the number of users grows, the value per coin increases. It has the potential for a positive feedback loop; as users increase, the value goes up, which could attract more users to take advantage of the increasing value" (Nakamoto, 2009).

2.3 Bitcoin As A Zero-to-One-Technology

Nakamoto's technological breakthrough does not lie in the invention of the individual components underlying Bitcoin's architecture. As we have shown above, technical and conceptual elements, such as hashing, proof-of-work, or time-stamping have existed before Nakamoto released the Bitcoin white paper (on 31 Oct. 2008), Genesis block (on 3 Jan. 2009) and code (on 9 Jan. 2009). Instead, the radical novelty of Bitcoin lies in how these technical components are organized and combined in the intricate design of the protocol, which gives rise to the singular system of an automated set of technological, economic, and social

[8] In contrast to earlier Internet protocols, such as TCP/IP or SMTP, which were difficult to monetize, Bitcoin directly motivates early adopters to adopt and hype the network.

incentives that have accelerated Bitcoin's rapid increase in adoption and value.[9] Bitcoin can thus be conceptualized as an assemblage of various technological components and its novelty arises from the radical re-combination of existing technologies. It instantiates an example of "combinatorial evolution," which complexity economist Brian Arthur has identified as the essential mode of technological innovation (see Arthur, 2009).[10] In other words, Nakamoto's invention represents, what venture capitalist Peter Thiel has termed, a vertical or "zero-to-one"-innovation. In contrast to horizontal innovations, which incrementally modify and improve existing technologies, zero-to one-innovations represent singularly radical new technologies (see Thiel and Masters, 2012).

Bitcoin was motivated by techno-libertarian ideals and cypherpunk beliefs in cryptographically-secured non-sovereign digital money, and therefore emerged in Internet sub-cultures (see Brunton, 2019). Given this origin, how could Bitcoin transition from a network with zero value and a single economic agent to an economic system consisting of several million users and valued at several billions dollars? How could Bitcoin become a zero-to-one technological breakthrough, with the transformative potential to disrupt the global monetary and financial system? In order to better understand how Bitcoin got bootstrapped and scaled, we need to dissect Bitcoin's underlying social dynamics and, in particular, understand the essential nature of bitcoin's bubbles. As we will show in the following sections, the bubbles that have punctuated the cryptocurrency's history do not simply instantiate a type of "collective hallucination" or "irrational exuberance" (see Odlyzko, 2010; Shiller, 2015). Rather, as we will argue, the design of the protocol itself already contains the seeds of these bubbles that have initiated the hype cycles so critical for the adoption of Bitcoin. They thus provide, as we will show, one of the purest examples of social bubbles in the sense of Gisler and Sornette (2009; 2010), Gisler et al. (2011), and Huber (2017).

3. The Prophecy of Satoshi Nakamoto: Religion, Hype, and Technological Adoption

[9] Contrary to the view that Bitcoin's underlying blockchain is the true technological innovation, we argue that the intertwined reflexive feedback loops that govern the protocol's design and incentive structure represent Bitcoin's novelty. Consequently, as it follows from our analysis, bitcoin, the cryptocurrency, cannot be separated from its underlying distributed ledger-technology as this would disrupt the intricate incentive system embedded in the network. Bitcoin commentator Joe Coin aptly captures the novelty of Bitcoin's design in a cogent blog post from 2015: "Given the crucial requirement to preserve decentralization, the problem Satoshi had to solve while designing Bitcoin was how to incentivize network participants to expend resources transmitting, validating, and storing transactions. The first step in solving that is the simple acknowledgement that it must provide them something of economic value in return […] The incentive had to be created and exist entirely within the network itself […] any instance of a blockchain and its underlying tokens are inextricably bound together. The token provides the fuel for the blockchain to operate, and the blockchain provides consensus on who owns which tokens. No amount of engineering can separate them" (Coin, 2015).

[10] An embryonic version of this idea can already be identified in Schumpeter's concept of "creative destruction." In 1910, he wrote that "to produce […] means to combine materials and forces within our reach [⋯]. To produce other things, or the same things by a different method, means to combine these materials and forces differently." Schumpeter identified creative destruction as a "source of energy within the economic system which would of itself disrupt any equilibrium that might be attained" (see Schumpeter, 1934). Creative destruction precisely refers to a combinatorial process in which novelty gets continually created by combining existing elements, which, in turn, constantly disrupt the established economic order.

As can be extracted from his communications on mailing lists, Satoshi Nakamoto started to code the protocol around May 2007. After he registered the domain bitcoin.org in August 2008, he started to send out emails drafts of the Bitcoin white paper. In October 2008, he publicly released the 9 pages long white paper that specifies the protocol and, soon after, released the initial code. On January 3, 2009, Nakamoto himself mined the so-called genesis block, that is, the first 50 bitcoins. In the same year, cryptographer and early Bitcoin-developer Hal Finney, which created the first reusable proof-of-work system before Bitcoin, received the first bitcoin transaction. By December 2010, others had taken over the maintenance of the project. On December 12, 2012, Nakamoto posted his final message to the Bitcoin forum.

As it has been noted by online commentators, the genesis of Bitcoin—its beginning as an obscure and radically novel technology invented by a mysterious pseudonymous creator that has disappeared—shares structural similarities with mythologies and religion. In this section, we will attempt to systematically deconstruct the analogies between Bitcoin and religious modes of social organization. We will argue that the quasi-religiosity of Bitcoin—that is, the fact that Bitcoin adopters are often labeled as “believers,” “evangelists,” or “cultists”—is a basic feature of Bitcoin’s technological diffusion process. In the next section, we will illuminate the structural analogy between the social dynamics governing Bitcoin adoption and religiosity. We will then outline the guiding visions that have emerged from the exegesis of Nakamoto’s whiter paper and conclude by examining in more detail Bitcoin’s model of technological diffusion, which is, to a large extent, based on memes and virality.

3.1 Bitcoin As Religion

Superficially, we can identify a few structural attributes of Bitcoin, which are analogous to religious history. As mentioned above, the most salient feature is the conceptual resemblance between Satoshi Nakamoto and religious or spiritual leaders, such as Abraham, Buddha, Jesus Christ, and their dedication towards their belief. Whereas Christ died by crucifixion as a sacrifice to achieve atonement for mankind’s sins, Nakamoto sacrificed his estimated 1,148,800 bitcoins that he has never moved from the original wallet (see Lerner, 2013). Similarly, the centrality of the white paper can be analogized to a sacred scripture in organized religions. The absence of Nakamoto—often referred to as Bitcoin’s “immaculate conception”—has stimulated competing exegeses of the white paper that aim to recover the true meaning of Nakamoto’s messianic vision of a decentralized digital form of money. Over the past decade, incompatible interpretations of the white paper relating to technical features, such as block-size limits, have triggered a series of hard-forks. Bitcoin Cash, for example, emerged in the summer of 2017 from developers’ disagreement about the block-sizes and transactions throughput. Consequently, different communities on Twitter, mailing lists, and online forums have organized around conflicting interpretations of the white paper and forks of the original Bitcoin source code, which represent the sacred object of Bitcoin. Culturally, the fragmentation into different “sects,” such as so-called “Bitcoin Maximalists”—which prioritize conservative protocol development and envision it as a settlement layer for large volume payments—or “Bcashers”—which emphasize Bitcoin as a payment system—has triggered many socio-cultural conflicts. As Bitcoin full-node operators choose which vision of Bitcoin they support by running the software that enforces the protocol rules, running nodes can be reinterpreted as one of the foundational ritual practices of Bitcoin. Analogous to religions, early disciples, such as technology entrepreneur Wences Casares, who spread Nakamoto’s utopian prophecy among Silicon

Valley venture capitalists, are a key-ingredient in the process of diffusing technological innovation. Naturally, for some of the more radical and dogmatic believers in the original vision of Nakamoto, the creation of so-called altcoins—that is, cryptocurrencies that either directly copy Bitcoin's source code or incorporate some of its technical or conceptual properties—is, in Bitcoin's eschatology, equalized to heresy[11]. Consequently, the heresy of attempting to clone Bitcoin's "immaculate conception" requires Bitcoin Maximalists to excommunicate altcoins and their developers and supporters from Bitcoin-related forums, social media platforms, and meetups.

While the conceptual similarity between Bitcoin and religion can be dismissed as irrelevant expressions of the social dynamics that govern Bitcoin sub-cultures, it is important to emphasize the critical importance of early adopters and their excessive commitment and enthusiasm for bootstrapping novel technologies. Before examining in more detail how the process of technology diffusion is unfolding in Bitcoin, we now briefly present an overview of the most dominant visions around which Bitcoin supporters have coalesced.

3.2 Satoshi's Vision

Not only did the timing of the release of the white paper and software coincide with the last great financial crisis, but the message embedded in the genesis block also contained a reference to the bank bailouts occurring in 2009. As we have alluded to above, Bitcoin's history reveals a genealogical link with various techno-libertarian and cypherpunk ideals around privacy and decentralization. Indeed, Nakamoto explicitly stated in an email to Hal Finney that Bitcoin is "very attractive to the libertarian viewpoint if we can explain it properly. I'm better with code than with words though" (see Nakamoto, 2008). In other words, Nakamoto literally encoded an ideological belief-system into the base layer of the protocol, which manifests itself in the decentralized and deflationary nature of Bitcoin. As his archived communications indicate—which are littered with references to central bank policies and failures of centralized modes of organization more generally—Nakamoto envisioned the protocol as a technological alternative to centralized economic systems. Bitcoin's design thus renders explicit its inherent normative role, which, in many other technologies, remains often elusive (see Radder, 2009). It is this intrinsic ideology that has catalyzed the early proselytization of Bitcoin, which, further promoted by the succession of bubbles that unfolded from 2012 to 2017 (Gerlach et al., 2019; Wheatley et al., 2019), accelerated its early adoption. While they cannot be cleanly separated, two foundational ideological views, which drive members of the Bitcoin community, can be broadly identified:

> *Bitcoin As Digital Gold*: This view, which is often inspired by Austrian Economics, emphasizes Bitcoin as a decentralized and "sound" alternative to fiat currency.[12] Given Bitcoin's finite and asymptotic money supply, supporters of this view—which, due to its monetary network effects,

[11] Interestingly, for theologist and philosopher René Girard, the original sin in Christianity lies in the "mimetic desire" of humans to imitate each other, which, ultimately, results in violence. On a Girardian reading, then, the emergence of altcoins, and the tribal rivalry and conflict these competing cryptocurrencies triggered, could be explained by the mimetic desire to copy the singularity of Bitcoin's design and successful implementation (see Hobart and Huber, 2019).
[12] For a discussion on "sound money" in Austrian Economics, see Ammous (2018).

consider Bitcoin to be the only legitimate cryptocurrency—believe that Bitcoin represents a digital substitute for gold. In this view, its hard-coded deflationary monetary policy and decentralized design, which enables censorship-resistance and reduces the risk of confiscation, makes Bitcoin a technologically more advanced store of value that is more secure than gold and state-issued fiat-currencies. As they envision it to compete with central banks and national fiat currencies and, thus, expect massive increases in value, most Bitcoin Maximalists are committed to hoarding bitcoins—or "hodling" as it is colloquially known. Consequently, for these Bitcoin Maximalists, the protocol's primary function is not to operate as a decentralized payment-network, which competes with centralized financial services, such as the SWIFT system or PayPal. Rather, they envision the Bitcoin network as a settlement layer in which block space is used to settle large value and high volume transactions—as opposed to facilitate individual small-value individual transactions (see Ammous, 2018). Instead, it is believed that Bitcoin needs to enable small and near-instantaneous transactions on a second layer, such as the Lightning Network, which is capable of settling millions of Lightning Network payments in one finalizing transaction on the Bitcoin-blockchain. Generally, Bitcoin Maximalists, and many Bitcoin Core developers, prefer a low-rate of innovation on the base layer and favor conservative protocol-development. While they envision a gradual ossification of the base layer, innovative experimentation is however encouraged on the second layer or so-called side-chains.

Bitcoin As Digital Cash: As mentioned above, Bitcoin Cash is the result of a hard-fork that occurred in mid-2017. Contrary to the Bitcoin Maximalist view that Bitcoin represents digitized gold, proponents of Bitcoin Cash generally cite the subtitle and abstract of the white paper and stress that Nakamoto's initial vision was to create a borderless, peer-to-peer electronic currency. In contrast to Bitcoin, Bitcoin Cash—which increased the block-size limit so that the network can process more transactions—is designed to establish itself first as medium of exchange and not as a store of value. For ideological and technological reasons, they favor on-chain activity and are opposed to the vision that the Bitcoin network should operate as a settlement layer due to fee increases. Therefore, based on the belief that payment activity will ensure its dominance, supporters of Bitcoin Cash encourage spending instead of "hodling." Bitcoin Core implemented Segregated Witness in a soft fork, which ensures a higher degree of decentralization as it enables users to run full nodes even on low-bandwidth connections. In contrast, Bitcoin Cash supporters believe that non-mining full nodes, which only receive and validate transactions, are not relevant to the security of the protocol. However, adoption of Bitcoin Cash has failed to materialize and on-chain activity eroded.

Flowing from these competing interpretations of Nakamoto's white paper are different visions of Bitcoin's future. A more moderate and pragmatic view holds that Bitcoin will adapt to regulatory constraints and integrate into the existing financial system. In this view, Bitcoin simply represents an uncorrelated asset class that can be used, similar to gold, to diversify and hedge portfolios against macroeconomic volatility and financial crises. Similarly, Bitcoin—conceived as a peer-to-peer payment network—can instantiate a decentralized and borderless alternative to centralized incumbent institutions and legacy financial networks, which are vulnerable to single-points of failures.

However, more radical futurist visions can be identified that derive from the philosophical foundations of Nakamoto's protocol, which reveal a quasi-religious set of beliefs. Believers in Hyperbitcoinization, for example, believe that the large-scale adoption of Bitcoin will result in a future demonetization of fiat currencies (see Krawisz, 2015). This belief is based on Bitcoin's censorship-resistant properties, which could undermine existing sovereign regulatory and political structures, and on superior monetary characteristics. In this view, the Bitcoin-induced collapse of fiat-currency and corresponding hyper-valuation of bitcoin, which might be triggered by systemic instabilities, such as a massive global financial crisis, would have far-reaching geo-political ramifications. Another futuristic view envisions Bitcoin as a breakthrough in information theory (see Gilder, 2018). Bitcoin, it is assumed, could serve as platform for general-purpose computation. Others even compare the Bitcoin network to a collective self-organizing intelligence (see Greenhall, 2016) or a "new form of life."[13]

Now, disregarding the plausibility and probability of such futuristic scenarios of Bitcoin adoption (see Senner and Sornette, 2019, for a critical review), these beliefs convey the Messianic dimension of Nakamoto's writings, which, for many proponents, promise technological salvation. However, it is precisely this set of philosophical beliefs, based on the different interpretations that we have outlined above, which has culturally accelerated the adoption of Bitcoin. In the next section, we will thus briefly describe Bitcoin's technological adoption cycles before we analyze in more detail how speculative Bitcoin bubbles accelerate the process of technology diffusion.

3.3 Bitcoin Evangelism and Technology Diffusion

Bitcoin has followed a specific technology adoption cycle that has been modulated by different social forces. Since its invention, we can discern four distinct, albeit idealized, phases in Bitcoin's diffusion. These phases of the ongoing technological adoption cycle have corresponded to bitcoin's price-acceleration bursts. In other words, each burst in bitcoin's price attracted a new set of adopters and resulted in a more widespread diffusion of the technology.

Bitcoin's technology adoption cycle has been initiated by a small group of believers, which constitutes the first cohort of adopters (see also Boyapati, 2018). Their adoption of this bleeding edge technology, which was motivated by the techno-ideological reasons illuminated earlier, elicited a process of continuous

[13] For example, cryptographer Ralph Merkle, who invented Merkle Trees—a data structure that Bitcoin employs—compares the protocol to an organism: " […]. Bitcoin is the first example of a new form of life. It lives and breathes on the internet. It lives because it can pay people to keep it alive. It lives because it performs a useful service that people will pay it to perform. It lives because anyone, anywhere, can run a copy of its code. It lives because all the running copies are constantly talking to each other. It lives because if any one copy is corrupted it is discarded, quickly and without any fuss or muss. It lives because it is radically transparent: anyone can see its code and see exactly what it does. It can't be changed. It can't be argued with. It can't be tampered with. It can't be corrupted. It can't be stopped. It can't even be interrupted […]. But as long as there are people who want to use it, it's very hard to kill, or corrupt, or stop, or interrupt" (see Merkle, 2016).

experimentation, debugging, and testing that gradually stabilized and improved the Bitcoin Core software. The quasi-religious devotion and extreme enthusiasm of this cohort of adopters, which consisted mainly of cryptographers, cypherpunks, and developers, then infected a group of ideologically motivated technologists, investors, and technology entrepreneurs, who in turn started to evangelize Bitcoin. The gradual build-out of the Bitcoin infrastructure and the first rudimentary exchanges, such as the infamous Japan-based Mt. Gox exchange, which allowed the conversion from fiat currency into Bitcoin, attracted early retail investors, which define the third cohort of adopters. The liquidity that these early speculators provided resulted in the first large-scale Bitcoin bubble in 2013, which triggered an inflow of more capital and attention. The launch of regulated exchanges, such as GDAX or Bitstamp, and OTC brokers, such as Cumberland Mining, in turn initiated the ongoing institutionalization of Bitcoin and intensified its virality.[14] If we now map Bitcoin's diffusion onto the generic technology adoption cycle, the fourth phase started after the bear market that followed the burst of the 2013 bubble, which lasted from 2014 to 2016. This phase of adoption is marked by the entry of the "early majority" of retail and institutional investors. Accelerated by the formation of regulated futures markets, such as the CME and CBOE, and other exchange-traded products, the price of bitcoin increased to almost 20'000 USD in December 2017. The infrastructure, which has been build out during the last bubble, might in the future usher the "late majority" and "laggards" phase of the technology adoption cycle.[15] Given that speculative frenzies boosted Bitcoin's adoption, in the next section, we provide a more granular analysis of how bubbles have catalyzed the adoption of the cryptocurrency.

4. Bitcoin-Bubbles As Innovation-Accelerators

4.1 A Brief Overview of Bitcoin Bubbles

Bitcoin's history is punctuated with speculative bubbles (Gerlach et al., 2019). Since the inception of the first exchange-traded price in 2010, the technological diffusion of Bitcoin can be conceptualized as a series of boom-bust cycles of increasing intensity and magnitude. This sequence of super-exponentially accelerating price-increases that are followed by equally spectacular crashes seems to follow the path of the classic Gartner Hype Cycle, which is used as a generic representation of the different phases of technology adoption. These hype cycles, which we will analyze in more detail below, have been fueled by speculate

[14] Bitcoin's diffusion occurred primarily online on social media, mailing lists, and blog posts. For example, Bitcoin's infectiousness has spread with various "memes," which acts as a unit for carrying and transmitting Bitcoin-related ideas and symbols. An example of Bitcoin's mimetic model of technology diffusion is the "Hodl"-meme, which—resulting from a misspelling of the word "hold"—motivates bitcoin holders to resist the urge to sell in response to market fluctuations. Consequently, analysts at Barclays developed an epidemiological model of Bitcoin's diffusion that models bitcoin as a "virus" that "infects" the population adopting the cryptocurrency technology.

[15] In Perez's classic conceptual model of technology diffusion, this phase might correspond to what she identifies as the "turning point." In her model, each technological disruption is triggered by a financial bubble, which allocates excessive capital to emerging technologies. Perez has extracted a regular generic pattern of technology-diffusion from historical case studies. She identifies an "installation"-phase in which a bubble drives the installation of the new technology. This is followed by the collapse of the bubble or a crash, to which she refers to as the "turning point." After this transitional phase—which occurred, for example, after the first British railway mania in the 1840s, or, more recently, after the dotcom-bubble—a second phase is unleashed: the "deployment" phase, which diffuses the new technology across economies, industries and societies (see Perez, 2003).

bubbles that, in turn, have generated more widespread diffusion of the technology. Consequently, each cycle corresponds to the distinct phases of adoption highlighted in the previous section.

We can identify five bitcoin-bubbles (see Figure 2) (see also Wheatley et al., 2019). In 2011, bitcoin's price increased from 1 USD on April 14 to 28.90 USD on June 9. In the following year, the price increased from 4.80 USD, on May 10, to 13.20 USD on August 15. In 2013, from January 3 to April 09, the price of bitcoin increased from 13.40 USD to 230 USD. In the same year, bitcoin increased from USD 123.20 on October 7 to USD 1156.10 on December 4. After the price crashed at the end of 2013, the price slowly recovered over a period of two years. On March 25, 2017, bitcoin's price started to accelerate from 975.70 USD to 20.089 USD on December 17, 2017, which represents bitcoin's all-time high. As this pattern of recurring bitcoin-bubbles demonstrates, each crash or correction was followed by an even larger-bubble in absolute prices (but of similar and very large amplitudes when measured in relative price changes). Bitcoin's price during the aforementioned bubbles was largely correlated with an increase in liquidity and with the maturation of the infrastructure, which attracted new adopters, such as entrepreneurs or speculators. While it was exceedingly difficult to trade bitcoin during the first bubble—which were primarily acquired through mining—exchanging and securing bitcoin has become relatively easy during the bitcoin-bubble that peaked in December 2017.

Bitcoin's price can thus be characterized by a hierarchy of repeating and exponentially increasing bubbles (Gerlach et al., 2019). These bubbles represent phases of unsustainable accelerating phases of price corrections and rebounds, which are driven by self-reinforcing feedback loops of herding behavior (Sornette, 2017). While the collapse of prices, which follows the faster-than-exponential power law growth processes defining bubble regimes (Sornette and Cauwels, 2015), can be destabilizing and destructive, bubbles of this type need to be understood as a source of technological innovation. By attracting capital in excess to what would be justified by a rational cost-benefit analysis or by a standard discounted cash flow calculation, bubbles accelerate the development of emerging technologies and, as Bitcoin clearly demonstrates, technology adoption cycles. Capital flows in at the early stage, which leads to a first wave of price increases. Attracted by the prospect of extrapolated higher returns, more investors follow, which triggers a positive feedback mechanism that fuels spiraling growth. Bubbles, which have historically incubated major technological innovations and disruptions, share a central dynamic: the funding of these new technologies decouples from rational expectations of economic return and, correspondingly, result in a reduction of collective risk-aversion. Irrespective of quantifiable financial returns and economic values, bubbles mobilize the financial capital needed to develop new transformative technologies. Based on the observation that these bubble dynamics extend beyond financial markets to social systems, one of the authors has, in a series of detailed case studies, developed the Social Bubble Hypothesis (see Sornette, 2008; Gisler and Sornette, 2009).

Based on the insight that many large-scale technological, social, or political projects involve collective enthusiasm and over-optimism, which results in unrestrained investment and commitments as well as a general reduction of risk aversion, the Social Bubble Hypothesis provides a useful conceptual framework to understand the emergence of Bitcoin.[16] In the next section, we will analyze the series of bitcoin bubbles identified above through the lens of the Social Bubble Hypothesis.

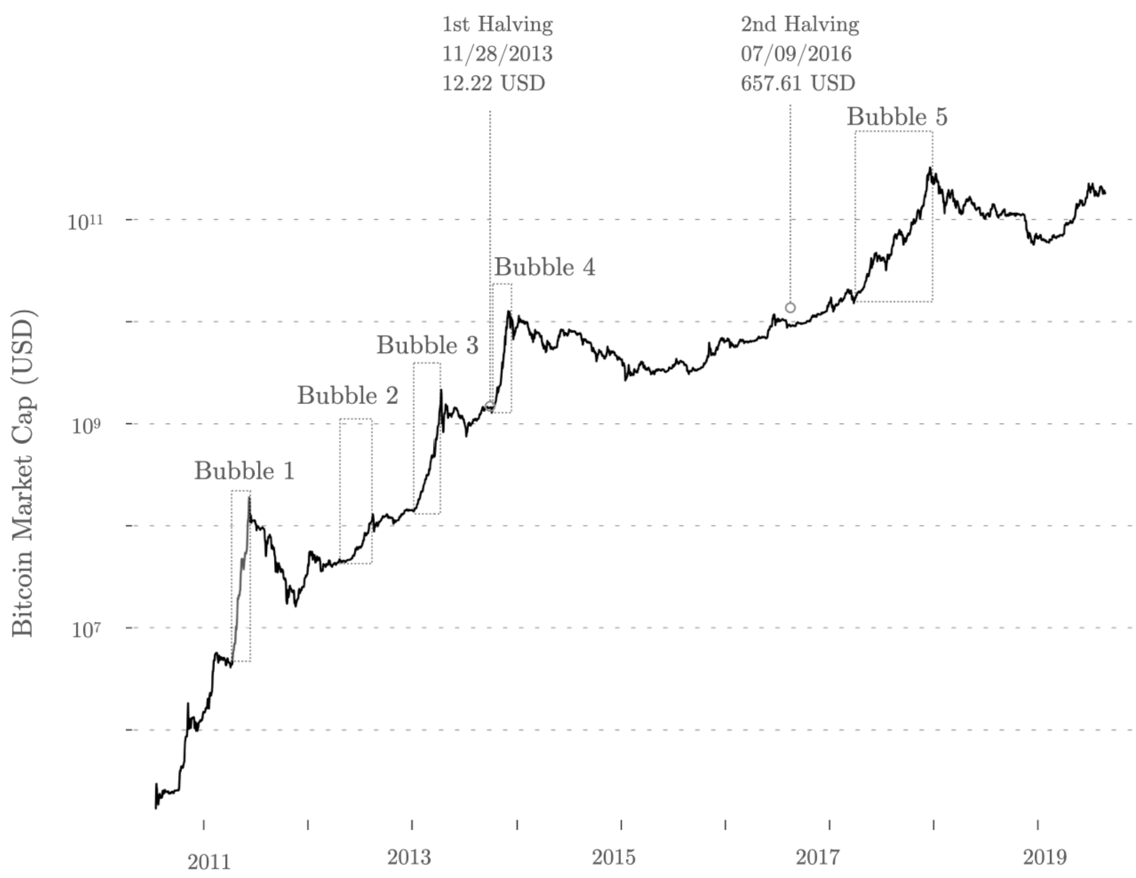


Figure 2: Main Bitcoin Bubbles and Hype Cycles. Notice the vertical scale and the larger than tenfold price increase in less than a year in each of these bubble episodes. Gerlach et al. (2019) document these bubbles as well as many other smaller ones covering this period.

4.2 Bitcoin As A Social Bubble

Similarly to financial bubbles, the essential ingredients of social bubbles are socio-behavioral mechanisms, such as herding or imitation, exuberant over-optimism and unrealistic expectations. By generating positive feedback cycles of extraordinary enthusiasm and investments—which have been essential for bootstrapping

[16] Mencius Moldbug, a pseudonym for technology entrepreneur and computer scientist Curtis Yarvin, embodies the spirit of the bubble dynamics of bitcoin, with his Bubble Theory of Money (BTM), which holds that, given its fundamentally social nature, money can be likened to a bubble. He writes: "Bitcoin is money and Bitcoin is a bubble, The BTM asserts that money and a bubble are the same thing. Both are anomalously overvalued assets. Both obtain their anomalous value from the fact that many people have bought the asset, without any intention to use it, but only to exchange it for some other asset at a later date. The two can be distinguished only in hindsight. If it popped, it was a bubble. If not, money—so far" (see Moldbug, 2013; Law, 2006). This reasoning should be distinguished from the standard theory of money, which considers it as an IOU and thus as credit (von Becke and Sornette, 2017), and the fact that credit growth is unstable and has led historically to boom-bust cycles over the last 5000 years (Graeber, 2012). In the creation of bitcoins, there is indeed no credit mechanism. The BTM is also reminiscent of the theory of value considered as a convention, developed by the French economist André Orléan (1987; 1989).

various social and technological enterprises—speculative bubbles can accelerate technological innovation. The complex networks of social interactions between enthusiastic supporters have catalyzed the formation of many large-scale scientific or technological projects. Characterized as manifestations of collective over-enthusiasm, they constitute an important element in the dynamics that give rise to scientific discoveries and radical technological breakthroughs. Similar to the generic technology hype cycle discussed previously, social bubbles are initiated by a burst of enthusiasm for a new technology. The earliest adopters and investors have strong convictions about the transformative nature of the technology they are investing in. The unbridled enthusiasm and commitments result in accelerated prices, which, in turn, catalyze more investments and speculation. Eventually, enthusiasm and investments peak, and the cycle is exhausted and prices and commitment saturate or decrease.

In the case of Bitcoin, in the early phase of the social bubble, the over-enthusiasm, commitment, and strong social interactions of cryptographers, computer scientists and cypherpunks significantly fueled the development and adoption of the technology. The enthusiasm and commitment of the cohort of early Bitcoin adopters then triggered the interest of early speculator and investors, which were often ideologically motivated to invest in the technology. It was this flow of capital and interest that triggered the first bitcoin bubbles in 2012 and 2013. After the peak of the first large bitcoin bubble, when bitcoin reached for the first time a price of more than 1000 USD in November 2013, the bubble collapsed and interest decreased substantially. The speculative fervor, which gave rise to super-exponential price growth, was then followed by despair, public derision, and a sense that the technology was not transformative at all. Eventually, bitcoin's price bottomed and went through a plateau during which a cohort of new believers and investors became attracted by the importance of the technology. Bitcoin's price-plateau persisted for two years before a new bubble gradually started to form in 2015. Over the prolonged bear market that lasted from 2013 to 2015, a new base of adopters has formed for the next iteration of the bubble cycle. This next iteration of the bubble, which in 2017 resulted in unprecedented hype and attention, attracted a much larger number of adopters. The cycle of bitcoin bubbles, which has given rise to accelerating prices and increasing media attention, has woven a network of reinforcing feedback loops that have led to widespread over-enthusiasm and commitment among Bitcoin Core developers, entrepreneurs, or speculators. This has been fueled by excessive expectations of ever-increasing price-acceleration and technology adoption. This momentous enthusiasm, which for instance led early cypherpunks and technologists to test and improve Bitcoin's code, and the extremely high expectations and hype towards the transformative potential of Bitcoin, constitute essential elements in the dynamics of Bitcoin's development and diffusion. In the next section, we will examine the nature of Bitcoin's sequence of bubble-driven hype cycles in more detail.

4.3 A Hierarchy of Bitcoin Hype Cycles, Speculative Bubbles, and Technological Adoption

Bitcoin's technological adoption—which could also reflect the monetization process of Bitcoin (see Boyapati, 2018)—seems to follow a hierarchical pattern of speculative bubbles within speculative bubbles that matches the shape of the classic Gartner hype cycle. The hype cycle, which represents the adoption of emerging transformative technologies, distinguishes between five phases. In the first phase, a technological breakthrough triggers initial interest. This phase corresponds to Nakamoto's release of the Bitcoin white paper and software, which attracted technologists and cypherpunks, such as Hal Finney who started to experiment with the Bitcoin technology when it was still in its proof-of-concept stage. Early adopters then started to improve the Bitcoin software. The first spike in Bitcoin's price occurred on July 12, 2010 on the first bitcoin exchange, The Bitcoin Market, after an article about Bitcoin Version 0.3 appeared the day before on the popular technology website site *Slashdot*. Following the launch of the Mt. Gox exchange in July 2010, bitcoin price peaked in June 2011 at 31.90 USD. During this first phase, Bitcoin entered the "Peak of Inflated Expectations" on the hype cycle. However, as the price of bitcoin decreased by over 93% over the following four months, Bitcoin entered the "Trough of Disillusionment," which is characterized by decreasing interest. After the price bottomed in April 2013, another price spike passed the psychological resistance level of 100 USD, which was fueled by the financial crisis in Cyprus that boosted bitcoin demand due to the growing distrust of banking from the threat of confiscation of banking deposits. However, after reaching a new high of 266 USD on Mt. Gox, it soon crashed below 60 USD before slowly returning to the range of 120 USD. After a longer phase of price stabilization, speculative investment resumed and Bitcoin entered the "Slope of Enlightenment," in which entrepreneurs have launched new Bitcoin-related startups and products. The next phase of technology adoption, which Bitcoin has not entered yet, is the "Plateau of Productivity" that is characterized by large-scale mainstream adoption.

While the trajectory of bitcoin's price since its inception can be mapped onto a generic hype cycle, it is important to note that each speculative bubble itself follows the path of a hype cycle. In other words, Bitcoin's technological adoption can be conceptualized as series of nested hype cycles, with a hierarchy of magnitudes and time scales. Unlike the generic Gartner hype cycle, however, Bitcoin's volatile curve of adoption does not follow a steady gradual increase. Instead, as Bitcoin's adoption has been speculative in nature, it followed a sequence of even more extreme growth phases than the standard exponentially growth path of the S-curve, which ended in a series of spectacular crashes. Generically, the hierarchical pattern of Bitcoin hype cycles seem to result from herding and imitation behavior of traders, which gives rise to speculative bubbles. As price accelerates, more speculators start to buy bitcoin. Eventually, as prices increase even more, early speculators are driven to take profits, which then triggers a correction or crash. Consequently, after each bubble-crash sequence, in which new long-term investors, or so-called "hodlers," are attracted, the amount of long-term holders increases during the bubble component of the cycle. In other words, due to its speculative bubbles, Bitcoin has been able to continually expand its adopter base.

Bitcoin's bubble-fueled technological adoption cycles can thus be conceptualized as a pattern of nested curves that each represent a new cohort of "hodlers." These subsequent waves of new "hodlers"—which represent future speculators who are not willing to sell in the next crash—can be quantified with a Bitcoin-native accounting structure called an UTXO—an "Unspent Transaction Output" (see Bansal, 2018). UTXO's, which are time-stamped by the transaction/block in which they were created, represent when a bitcoin was last used in a transaction. We can identify different adoption waves since Bitcoin's release, which occur when a cohort of new speculators or investors buy bitcoins during a bubble and hold through the downturn into the next market cycle. Visually, in Figure 4, these speculative adoption waves can be represented by different age bands: whereas warmer-colored age bands (<1 day, 1 day–1 week, 1 week–1 month) represent transactions of large amounts of bitcoin, the steady growth of the top, cooler-colored age bands (2–3 years, 3–5 years, >5 years) indicate the adoption of Bitcoin, that is, they represent an increase in “hodling.”[17] These adoption waves manifest themselves visually as nested curves, which are caused by each age band becoming progressively wider (see Figure 4). The different levels of unspent transaction outputs indicate that each bubble attracts a new cohort of “hodlers” who are accumulating and holding bitcoin. In other words, each speculative bubble has triggered a “hodling” wave. Bitcoin thus represents one of the purest examples of *speculative technological adoption*. The bubble-driven repeating and super-exponentially increasing hype cycles continually attracted new cohorts of “hodlers.

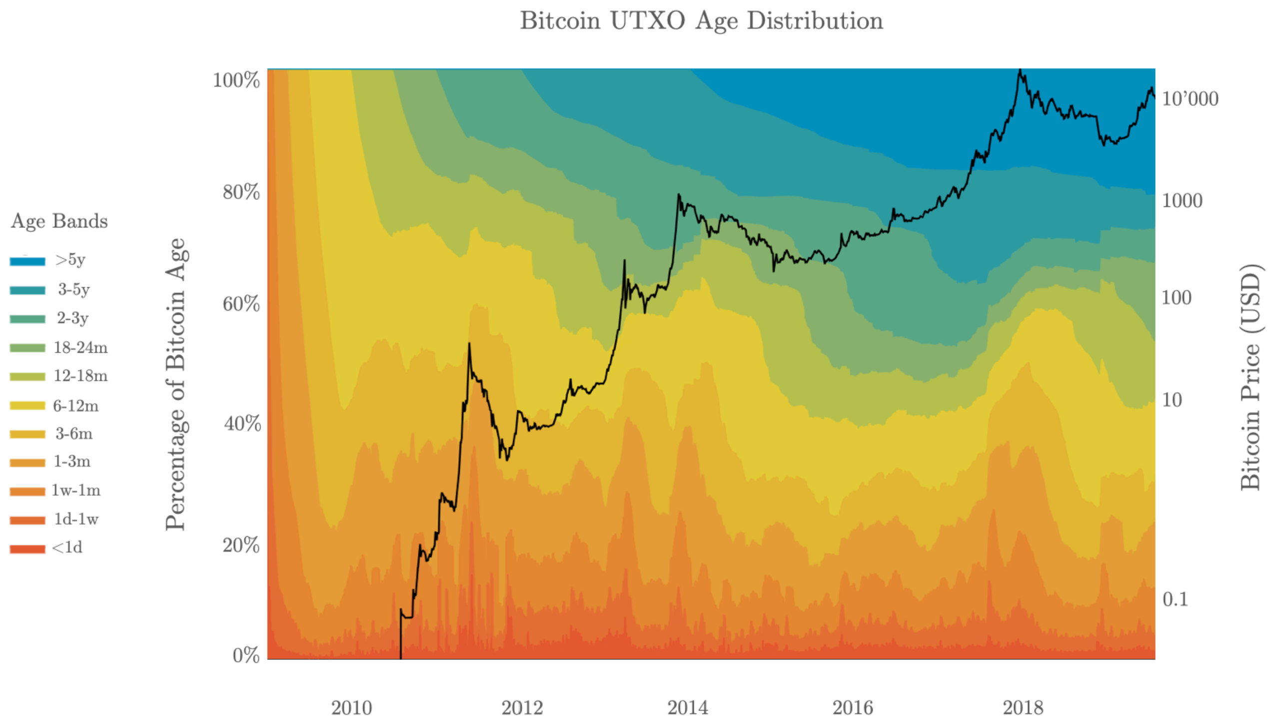


Figure 3: Speculative Bitcoin Adoption Waves in colors (left axis and color codes) superimposed on the Bitcoin price (black line and right axis) (adopted from Bansal, 2018).

4.4 Bitcoin’s Technological Adoption Cycle

It is interesting to note that the hype-cycles fueling Bitcoin’s technological adoption are embedded in the protocol itself. In Nakamoto’s design, as mentioned above, every four years a halving occurs that reduces the reward for miners by half. Built into the protocol to control Bitcoin’s inflation, the previous so-called halvening have coincided with massive price-accelerations. After the first halvening, which occurred in November 2012, Bitcoin’s price increased from 12 USD to more than 650 USD at the time when the second halvening in July 2016 occurred. After the block reward reduction to 12.5 bitcoins, where each block is created every 10 minutes, the price accelerated to almost 20’000 USD. While it is of course uncertain whether the next halvening —which will occur in mid-2020—will accelerate prices in a similar way, the previous halvening have fueled the main Bitcoin’s hype cycles. As his message mentioned above indicates, Nakamoto seemed to have programmed speculative bubbles into the protocol, with the intention of accelerating the feedback loops needed to bootstrap bitcoin’s value. In the message posted on the P2P Forum in 2009 Nakamoto stated: “As the number of users grows, the value per coin increases. It has the potential

[17] “Hodling” represents the speculative adoption of bitcoin as it implies a speculative bet on future gains in value. Thus, adoption is here defined as an increase in “hodling”—that is, the accumulation and holding of bitcoin.

for a positive feedback loop; as users increase, the value goes up, which could attract more users to take advantage of the increasing value" (Nakomato, 2009). Positive feedback is known to be the main mechanism for the generation of bubbles (Sornette, 2017; Johansen and Sornette, 2010; Jiang et al., 2010; Sornette and Cauwels, 2015). It is thus plausible that Nakamoto designed the halvenings to create "artificial" boom-and-bust cycles. The 4-year-halvening cycles drive up prices, which are then followed by an increase in hash rate and number of hodlers. Even after a crash, the hash rate and, consequently, the security of the network are higher than before the price-acceleration. Moreover, the halvening also attracts new cohorts of adopters to "hodl."[18] Over the past decade, this sequence of speculative bubbles thus bootstrapped a new form of digital money, which started with zero value (and arguably with zero fundamental value in the standard economic sense) to a network that, at the peak of the last bubble, was valued at more than USD 320 billion. Rather than simply representing excessive speculation, we can conclude that bubbles and hype cycles have been accelerating Bitcoin's technological adoption. In other words, speculative bubbles provide the fundamental mechanism fueling the intertwined techno-economic feedback loops that drive Bitcoin's security, value, and network effects.

5. Conclusion and Discussion

We have elaborated on the elements supporting the claim that Bitcoin represents a technological breakthrough. As we have noted, Bitcoin represents a radical technological innovation, not simply because it represents a novel technology. Rather, the novelty emerges from Nakamoto's combination of ideas and technologies that existed previously in disparate and previously unrelated fields. As its historical genesis shows, Bitcoin represents an instance of "radical" or "vertical" technological innovation—its "novelty" emerged from an assemblage of existing technological components. The invention of Bitcoin thus required the bridging of disparate fields, terminologies, and assumptions. Bitcoin's breakthrough lies in how Nakamoto designed a system of interlocking techno-economic feedback loops that fuel its value, security, and network effects.

Bitcoin genesis is different from any technological breakthrough that has historically preceded it. As we have documented above, Bitcoin's "immaculate conception" by a pseudonymous programmer has given rise to a community of developers and users who have a quasi-religious commitment to the cryptocurrency. Given Bitcoin's open-sourced design and distributed architecture, the technology represents a permissionless and decentralized model of innovation. Bitcoin did not result—as it was historically the case with preceding technological breakthroughs—from a specific set of innovation policies or government-funded academic research. Rather, it emerged outside the boundaries of academic peer-review or government-funding. It

[18] The expectation that the halvening results in an increase of "hodlers" and higher prices assumes that scarcity drives bitcoin's value. In this view, the halvening represents a supply "shock" that increases bitcoin's relative scarcity. While the supply cap of 21 million bitcoins is algorithmically fixed, relative supply and the bitcoins in circulation decrease. Furthermore, due to the change in the supply schedule, the market needs to absorb fewer bitcoins, which miners are selling to cover their capital expenditures. Gradually, miner compensation will transition to transaction fees. A popular model that is used to model bitcoin's scarcity-based value is the so-called Stock-to-Flow model. It models the price of Bitcoin based on the "stock-to-flow ratio," which was initially used to value gold and other raw materials. By relating the "stock"—i.e., the quantity issued–to the "flow"—i.e., the annual issued quantity–the model derives a prediction of a bitcoin price post-halvening of $55,000 to $100'000 (which would correspond to a market cap of more than $1 trillion) (see Plan B, 2019).

appeared on an obscure mailing list, which was adopted and diffused by a group of extremely committed and enthusiastic supporters. In this sense, Bitcoin and the self-organizing principles that govern the protocol's evolution share an essential similarity with the emergent properties that complex systems exhibit, which Hayek characterized as "spontaneous order" (see Hayek, 1969). It will be interesting to observe whether the beliefs of these enthusiastic supporters, which evangelize Bitcoin, will in the future follow the trajectory of a self-fulfilling prophecy that will continually attract new developers, entrepreneurs, and "hodlers" as it has been the case until now. Furthermore, as we have shown above, what is singularly unique in Bitcoin is that hype-cycles are built into the design of the protocol itself. The deflationary nature of the cryptocurrency's supply and halving of block-rewards have triggered a process of what can be called *speculative technology adoption*. The resulting hierarchical sequence of repeating and super-exponentially increasing series of bubbles, which have occurred over the past decade since Bitcoin's inception, has resulted in new waves of speculative adopters.

As we have argued, these speculative bubbles and hype-cycles have bootstrapped the Bitcoin network into existence. Whereas financial bubbles have historically been important catalysts in the diffusion of technological revolutions, Bitcoin represents the first radical technological innovation in which bubbles constitute necessary components in the process of technology adoption and diffusion. As our previous discussion of the halvening-cycle has shown, the emergence of bubbles and hype-cycles, which accelerate the technological adoption, are hard-coded into the protocol. Yet, the fundamental question remains of whether Bitcoin's technological breakthrough can be replicated. Can Bitcoin's "immaculate conception"—that is, its invention by a pseudonymous programmer, which attracted a following of committed believers—get repeated? While Bitcoin's invention represents a technological singularity,[19] its history nevertheless demonstrates the importance of hype and bubbles for the development and diffusion of cutting-edge technologies. A more generalizable insight, which can be derived from our Bitcoin case study, is that bubbles need to be conceptualized as "chaotic attractors" for technological innovation. The development of a definite vision of the future—which generates hype and great enthusiasm that reduces risk-aversion and attracts new supporters and adopters—is driving the Bitcoin phenomenon. While the Bitcoin experiment is still unfolding, its emergence has clearly demonstrated that hype and bubbles constitute essential elements in the process of technological innovation. In order to be successful, future technological innovations, we suggest, will thus need to incubate and generate comparable visions, enthusiasm, and hype, which are currently fueling the dynamics of Bitcoin's development and diffusion.

[19] Bitcoin could only be invented once. Given its singular nature, it has been argued that Bitcoin cannot be replaced by another cryptocurrency. Hal Finney, for example, stated that every subsequent version of the protocol designed to substitute Bitcoin would be self-invalidating. A hypothetical Bitcoin successor would undermine its own viability and credibility as "an investor" would not "know that it won't happen again" (Finney, 2011). In this view, the adoption of Bitcoin follows a binary logic: either Bitcoin succeeds or Bitcoin and all other forks or cryptocurrencies will fail as well.

References:

Abreu, D. and M.K. Brunnermeier, Bubbles and Crashes, *Econometrica* 71.1, 173-204 (2003).

Allen, F. and G. Gorton, Churning Bubbles, *The Review of Economic Studies* 60.4, 813-836 (1993).

Ammous, S., *The Bitcoin Standard: The Decentralized Alternative to Central Banking*. John Wiley & Sons (2018).

Antonopoulos, Andreas M., *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*. O'Reilly Media (2014).

Arthur, W. B., *The Nature of Technology: What It Is and How It Evolves*. Simon and Schuster (2009).

Bansal, D., Bitcoin Data Science (Pt. 1): HODL Waves, (https://unchained-capital.com/blog/hodl-waves-1/) (2018).

Boyapati, V., The Bullish Case for Bitcoin, (https://medium.com/@vijayboyapati/the-bullish-case-for bitcoin-6ecc8bdecc1) (2018).

Brunton, F., *Digital Cash: The Unknown History of the Anarchists, Utopians, and Technologists Who Created Cryptocurrency*. Princeton University Press (2019).

Casey, M.B., Speculative Bitcoin Adoption/Price Theory, (https://medium.com/@mcasey0827/speculative bitcoin-adoption-price-theory-2eed48ecf7da) (2016).

Chaum, D., Blind Signatures For Untraceable Payments, *Advances in Cryptology Proceedings of Crypto* 82 (3): 199–20 (1983).

Chohan, U.W., A History of Bitcoin, (https://ssrn.com/abstract=3047875) (2017).

Coin, J., Crypto 2.0 and Other Misconceptions, (https://medium.com/@joecoin/crypto-2-0-and-other misconceptions-6bbe9c002cf6) (2015).

Dai, W., B-money, (http://www.weidai.com/bmoney.txt) (1998).

Dwork, C., and M. Naor, Pricing via Processing or Combatting Junk Mail, *Annual International Cryptology Conference*. Springer (1992).

Finney, H., Early Speculators' Reward, (https://bitcointalk.org/index.php?topic=10666.0) (2011)

Garber, P.M., *Famous First Bubbles: The Fundamentals of Early Manias*. MIT Press (2001).

Gerlach, J.-C., G. Demos and D. Sornette, Dissection of Bitcoin's Multiscale Bubble History from January 2012 to February 2018, *Royal Society Open Science* 6, 180643 (2019).

Gilder, G., *Life after Google: The Fall of Big Data and the Rise of the Blockchain Economy*. Simon and Schuster (2018).

Gisler, M., and D. Sornette., Exuberant Innovations: The Apollo Program, *Society* 46.1, 55-68 (2009).

Gisler, M., Sornette D., and R. Woodard, Innovation As a Social Bubble: The Example of the Human Genome Project, *Research Policy* 40.10, 1412-1425 (2011).

Graeber, D., *Debt: The First 5,000 Years*, Melville House; reprint edition (2012).

Greenhall, J., Lets Get This Straight, Bitcoin is an Experiment in Self-Organizing Collective Intelligence, (https://blog.p2pfoundation.net/lets-get-straight-bitcoin-experiment-self-organizing collectiveintelligence/2017/05/16) (2016).

Haber, S. and W. Stornetta, How to Time-Stamp a Digital Document, *Crypto'90*, LNCS 537, 437-455 (1991).

Hayek, F.A., *Studies in Philosophy, Politics and Economics*. Touchstone (1969).

Hobart, B. and T. Huber, Manias and Mimesis: Applying René Girard's Mimetic Theory to Financial Bubbles, (https://dx.doi.org/10.2139/ssrn.3469465) (2019).

Huber T., Bubbles As Innovation Accelerators, *NewCo Shift*, (https://medium.com/newco/innovative-exuberance-ad75ee39f4c5) (2017).

Jakobsson, M. and A. Juels, Proofs of Work and Bread Pudding Protocols, *Secure Information Networks: Communications and Multimedia Security*. Kluwer Academic Publishers: 258–272 (1999).

Janeway, W. H., *Doing Capitalism in the Innovation Economy*. Cambridge University Press (2012).

Jiang, Z.-Q., W-X. Zhou, D. Sornette, R. Woodard, K. Bastiaensen and P. Cauwels, Bubble Diagnosis and Prediction of the 2005-2007 and 2008-2009 Chinese Stock Market Bubbles, *Journal of Economic Behavior and Organization* 74, 149-162 (2010).

Johansen, A. and D. Sornette, Shocks, Crashes and Bubbles in Financial Markets, *Brussels Economic Review* (Cahiers economiques de Bruxelles) 53 (2), 201-253 (2010).

Krawisz, D., Hyperbitcoinization, (https://nakamotoinstitute.org/mempool/hyperbitcoinization/) (2014)

Land, N., *Crypto-Current*, (http://www.uf-blog.net/) (forthcoming).

Law, J., Why the Global Financial System is about to Collapse, (https://safehaven.com/article/5165/why the-global-financial-system-is-about-to-collapse) (2006).

Lerner, S.D, The Well Deserved Fortune of Satoshi Nakamoto, Bitcoin creator, Visionary and Genius, (https://bitslog.com/2013/04/17/the-well-deserved-fortune-of-satoshi-nakamoto/) (2013).

Merkle, R., DAOs, Democracy and Governance, (http://merkle.com/papers/DAOdemocracyDraft.pdf) (2016).

Merton, R.K., The Self-Fulfilling Prophecy, *The Antioch Review* 8.2, 193-210 (1948).

Moldbug, M., Bitcoin is Money, Bitcoin is a Bubble, (https://www.unqualified-reservations.org/2013/04/bitcoin-is-money-bitcoin-is-bubble) (2013).

Nakamoto, S., Bitcoin Open Source Implementation of P2P Currency, (https://satoshi.nakamotoinstitute.org/posts/p2pfoundation/2/) (2009).

Nakamoto, S., Bitcoin: A Peer-to-Peer Electronic Cash System, (https://bitcoin.org/bitcoin.pdf) (2008).

Narayanan, A. and J. Clark, Bitcoin's Academic Pedigree. *Communications of the ACM*, 60(12), 36-45 (2017).

Narayanan, A., Bonneau, J., Felten, E., Miller, A. and S. Goldfeder, *Bitcoin and Cryptocurrencytechnologies: A Comprehensive Introduction*. Princeton University Press (2016).

Odlyzko, A., *Collective Hallucinations and Inefficient Markets: The British Railway Mania of the 1840s*. University of Minnesota (2010).

Orléan, A., Anticipations et conventions en situation d'incertitude, *Cahiers d'Économie Politique* 13, 153-172 (1987).

Orléan, A., Pour une approche cognitive des conventions économiques, *Revue Economique* 40 (2), 241-272 (1989)

Perez, C., *Technological Revolutions and Financial Capital*. Edward Elgar Publishing (2003).

Plan B, Modeling Bitcoin's Value with Scarcity, (https://medium.com/@100trillionUSD/modeling-bitcoins-value-with-scarcity-91fa0fc03e25) (2019).

Radder, H., Why Technologies Are Inherently Normative, *Philosophy of Technology and Engineering Sciences*, 887-921, North-Holland (2009).

Rochard P., Bitcoin Governance, (https://medium.com/@pierre_rochard/bitcoin-governance-37e86299470f) (2018).

Santos, M.S., and M. Woodford, Rational Asset Pricing Bubbles, *Econometrica: Journal of the Econometric Society*, 19-57 (1997).

Scheinkman, J.A. and W. Xiong, Overconfidence and Speculative Bubbles, *Journal of Political Economy* 11.6, 1183-1220 (2003).

Schumpeter, J.A., *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*, Harvard University Press (1934).

Senner, R. and D. Sornette, The Holy Grail of Crypto Currencies: Ready to Replace Fiat Money? *Journal of Economic Issues* 53 (4), 966-1000 (December 2019) (http://ssrn.com/abstract=3192924).

Shiller, R.J., *Irrational Exuberance*. Princeton University Press, revised and expanded third edition (2015).

Song, Jimmy. *Programming Bitcoin: Learn How to Program Bitcoin from Scratch*, O'Reilly Media (2019).

Sornette, D. and P. Cauwels, Financial Bubbles: Mechanisms and Diagnostics, *Review of Behavioral Economics* 2 (3), 279-305 (2015).

Sornette, D., *Why Stock Markets Crash (Critical Events in Complex Financial Systems)*, Princeton University Press, revised edition (2017).

Sornette, Didier. Nurturing Breakthroughs: Lessons from Complexity Theory. *Journal of Economic Interaction and Coordination* 3.2 (2008): 165.

Szabo, N.. Bit gold, (https://unenumerated.blogspot.com/2005/12/bit-gold.html) (2008).

Thiel, P.A. and B. Masters. *Zero to One: Notes on Startups, or How to Build the Future*. Broadway Business, (2014).

von der Becke, S. and D. Sornette, Should Banks be Banned from Creating Money? An Analysis from the Perspective of Hierarchical Money, *Journal of Economic* Issues 51 (4), 1019-1032 (2017).

Wheatley, S., D. Sornette, M. Reppen, T. Huber and R.N. Gantner, Are Bitcoin Bubbles Predictable? Combining a Generalised Metcalfe's Law and the LPPLS Model, *Royal Society Open Science* 6, 180538, (http://dx.doi.org/10.1098/rsos.18053), 1-13 (2019).

Swiss Finance Institute

Swiss Finance Institute (SFI) is the national center for fundamental research, doctoral training, knowledge exchange, and continuing education in the fields of banking and finance. SFI's mission is to grow knowledge capital for the Swiss financial marketplace. Created in 2006 as a public–private partnership, SFI is a common initiative of the Swiss finance industry, leading Swiss universities, and the Swiss Confederation.

swiss:finance:institute

c/o University of Geneva, Bd. Du Pont d'Arve 42, CH-1211 Geneva 4
T +41 22 379 84 71, rps@sfi.ch, www.sfi.ch

4. Social Bubbles: Clean-Tech Bubble Case Study

Full Reference:

Gregoris V., Huber, T. and Sornette, D. (2021): "Salvation and Profit": Deconstructing the Clean-Tech Bubble. *Swiss Finance Institute Research Paper Series* Nº 21-36.

Abstract

From 2004 to 2008, a bubble formed in clean technologies, such as solar, biofuels, batteries, and other renewable energy sources. In this paper, we analyze this clean-tech bubble through the lens of the Social Bubble Hypothesis, which holds that strong social interactions between enthusiastic supporters weave a network of reinforcing feedbacks that lead to widespread endorsement and extraordinary commitment by those involved. We present a detailed synthesis of the development of the clean-tech bubble, its history, and the role of venture capital and government funding in catalyzing it. In particular, we dissect the underlying narrative that was fueling the bubble. As bubbles can be essential in the process of accelerating the development of emerging technologies and diffusion of technological innovations, we present evidence that the clean-tech bubble constituted an example of an innovation-accelerating process.

RESEARCH

Swiss Finance Institute Research Paper Series N°21-36

"Salvation and Profit":
Deconstructing the Clean-Tech Bubble

Vincent Giorgis
ETH Zurich

Tobias A. Huber
ETH Zurich

Didier Sornette
ETH Zurich, Southern University of Science and Technology (SUSTech) Shenzhen, Tokyo Institute of Technology, and Swiss Finance Institute

“Salvation and Profit”:
Deconstructing the Clean-Tech Bubble

Vincent Giorgis[1], Tobias A. Huber[1] and Didier Sornette[1,2]

[1]ETH Zurich, Department of Management, Technology and Economics, Switzerland
[2] Institute of Risk Analysis, Prediction and Management (Risks-X), Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology (SUSTech), Shenzhen, 518055, China

May 23, 2021

Abstract: From 2004 to 2008, a bubble formed in clean technologies, such as solar, biofuels, batteries, and other renewable energy sources. In this paper, we analyze this clean-tech bubble through the lens of the Social Bubble Hypothesis, which holds that strong social interactions between enthusiastic supporters weave a network of reinforcing feedbacks that lead to widespread endorsement and extraordinary commitment by those involved. We present a detailed synthesis of the development of the clean-tech bubble, its history, and the role of venture capital and government funding in catalyzing it. In particular, we dissect the underlying narrative that was fueling the bubble. As bubbles can be essential in the process of accelerating the development of emerging technologies and diffusion of technological innovations, we present evidence that the clean-tech bubble constituted an example of an innovation-accelerating process.

Keywords: *Financial Bubbles, Narrative Economics, Technological Innovation, Clean Tech, Energy, Venture Capital*

JEL: C54, D61, D70, F64, G01, O25

1. Introduction

On March 8, 2007, legendary venture capitalist John Doerr, who has invested in Netscape, Amazon, and Google, gave a talk entitled “Salvation (and Profits) in Greentech.” In his talk, Doerr described climate change as “the largest economic opportunity of the 21 century, and a moral imperative” (Doer, 2007). The title of Doerr’s talk perfectly encapsulates the simultaneous belief in profits and salvation that fueled the clean-tech bubble, which, a year later, spectacularly crashed. As venture capitalist Peter Thiel stated: “Instead of a healthier planet, we got a massive clean tech bubble” (Thiel and Masters, 2012). In this paper, we analyze the clean-tech bubble and the economic, social, and political factors that catalyzed it. In particular, we address the question whether the clean-tech bubble developed the virtue of having accelerated innovation in clean and renewable energy technologies.

In order to tackle such questions, we draw on the Social Bubble Hypothesis, which holds that, under specific conditions and designs, speculative bubbles can be important components in the process of innovation. At the core of previous social bubbles, we have identified social interactions between enthusiastic supporters, which weave a network of reinforcing feedbacks that lead to widespread endorsement and extraordinary commitment by those involved (Sornette, 2008; Gisler and Sornette, 2009, 2011; Gisler et al., 2011). As we will show below, the clean-tech bubble followed a similar pattern as it was fueled by self-reinforcing feedback loops of commitment and excessive enthusiasm, which one technology investor has even compared to "religious thinking".

In the economics and finance literature, speculative financial bubbles form when unrealistic expectations about future cash flows decouple prices temporally from fundamental valuations. As unrealistic and excessively optimistic expectations about the future end up failing, bubbles are considered to have predominately destructive social and economic effects (Chauvin et al., 2011; Farmer, 2015). In standard finance and economics, bubbles, thus, are associated with market failures. In contrast, the Social Bubble Hypothesis was proposed to develop an alternative view of bubbles that illuminates the positive dimension of periods "popular delusions," which have regularly disrupted markets (see Garber, 2001; Kindelberger, 2017). Indeed, while most speculative bubbles distort prices and temporarily destabilize markets, the formation of certain bubbles can act as important catalysts for socio-technological innovation. Such bubbles, as they deploy the financial capital necessary to fund disruptive technologies at the frontier of innovation, are thus capable of accelerating breakthroughs in science, technology, and engineering. By generating positive feedback loops of—what appears in the short-term as—excessive enthusiasm or even malinvestments, they have been essential for bootstrapping various social and technological innovations (Gisler and Sornette, 2009, 2011; Gisler et al., 2011). As they accelerate capital spending cycles that stimulate over-investment in infrastructures or emerging technologies, some bubbles have in the past enhanced economic productivity. Historically, this dynamic has been, for example, identified in the build-out of foundational economic infrastructures, such as canals, railroads, telecom, or energy networks. In these examples, the build-up of excess capacity financed by speculative bubbles could still get exploited even after the collapse of the preceding speculative bubble has recalibrated investor expectations—often by a new set of market participants (for historical examples, see Nairn, 2002).

Characterized as manifestations of collective over-enthusiasm, these phases of "irrational exuberance" (Shiller, 2015) constitute an important element in the dynamics that give rise to great scientific discoveries and radical technological breakthroughs throughout history. It is therefore not surprising that financial bubbles and speculative manias have been at the core of the technological revolutions that have fundamentally transformed our economic, social, and technological systems over the last centuries (see Perez, 2002; Sornette, 2008; Janeway, 2012).[1] It is this complex networks of social and financial interactions between enthusiastic supporters and speculators that have catalyzed the formation of many, more recent large-scale scientific or technological projects, such as the Apollo Program, the Human Genome Project, or the development of Bitcoin (see Sornette, 2008, Gisler and Sornette, 2009; 2011; Gisler et al., 2010; Huber and Sornette, 2016; Huber and Sornette, 2020).

It thus seem natural to examine the clean-tech bubble in the framework of social bubbles. After the bust of the dotcom-bubble, this bubble formed as private and public investors started to envision clean-tech as the next frontier for technological innovation. Within the framework of social bubbles, the crucial question that needs to be addressed is to what degree the clean-tech bubble has really been accelerating innovation in clean technologies. Whereas venture capital (VC) investment in renewable energy measured USD 286 million in 2001, more than USD 4.1 billion flowed—at the peak of the bubble in 2008—into the sector. Fueled by volatile silicon prices, falling natural gas prices due to the widespread adoption of hydraulic fracking in the US, the 2008 financial crisis, and China's emerging solar industry, the bubble spectacularly crashed in 2008 and VC investments fell to USD 2.5 billion in 2009 (see Yergin, 2011). Now, in order to assess whether the clean-tech mania accelerated the development and adoption of clean technologies, we will examine in what follows the quantitative as well as qualitative factors that gave rise to the bubble.

The paper is organized as follows. The first section provides a brief history of the clean-tech bubble. We then present a quantitative analysis of the bubble. We analyze in more detail VC investment in clean-tech and identify the challenges that arise for investments in this sector. In the next section, we analyze the clean-tech bubble through the lens of the Social Bubble Hypothesis. Here, we especially focus on the role of the narratives driving the bubble. We conclude by providing in the last part of the paper a discussion of bubbles as innovation-accelerators and address the question whether the clean-tech bubble of the last decade catalyzed technological innovation and adoption of clean technologies.

2. The Clean-Tech Bubble: A Short History

After the collapse of the dotcom-bubble in 2000, which destroyed around USD 5 trillion in (previously excessively accumulated) market capitalization, in need for new investment horizons, venture capital investors started to envision clean-tech as the next frontier of technological innovation. For example, Vinod Khosla, co-founder of Sun Microsystems, started to invest with his VC firm Khosla Ventures in biofuels and other renewables. Clean-tech firms, such as the solar panel maker Energy Innovations, Nanosolar, or Recurrent Energy, were launched and Elon Musk, who previously co-founded fintech startup PayPal, invested USD 96 million in the electric car-startup Tesla Motors, alongside well-known VC Steve Jurvetson. In 2008, Kleiner Perkins, one of the most well-known VC firms in Silicon Vally, allocated more than USD 300 million to a clean-

[1] Joseph Schumpeter identified the role of bubbles in technological innovation already in the early 20th century. He writes about the South Sea Bubble of 1720: "The mania of 1719-20 […] was, exactly as were later manias of this kind, induced by a preceding period of innovation which transformed the structure and upset the preexisting state of things" (Schumpeter, 1939, p. 250).

tech fund and USD 500 million to a growth fund “intended to help speed mass-market adoption of solutions to the world’s climate crisis.” However, venture capital was not the only factor driving the build-up of the clean-tech bubble.

At the start of the 21st century, debates about climate change intensified, catalyzed by the occurrence of extreme weather events, such as a heat wave in Europe in the Summer of 2003 and Hurricane Katarina in 2005, and influenced by the pervading presence of the media (Vasterman et al., 2005) and later by the exploding growth of social media (Peng, 2013; Aral, 2020). The offshore oil industry was devastated by Katrina, which reduced the oil supply and resulted in a rising oil price. The consequences of these natural disasters, the release of new scientific studies, such as a report of the Intergovernmental Panel on Climate Change (IPCC) on the anthropogenic nature of climate change in 2007, and the release in 2006 of Al Gore’s documentary, “An Inconvenient Truth”, started to fuel interest in renewable and clean energy as an alternative to the status quo. Investing and developing green energy sources was becoming for many a moral imperative “to save the Planet”.

Not only was there an increasing inflow of venture capital, but the federal government started also to invest in emerging clean technologies. In 2005, as part of the Energy Policy Act, a federal loan guarantee program was authorized at USD 4 billion “to support innovative clean energy technologies that are typically unable to obtain conventional private financing due to high technology risks.” In total, the loan guarantee program provided more than USD 16 billion for 28 clean-tech projects. For example, Solyndra, a prominent clean-tech start-up that manufactured cylindrical solar tubes, had received USD 500 million in federal loan guarantees, but filed later for bankruptcy. Electric vehicle startups Tesla and Fisker received USD 465 million and USD 539 million to open factories, respectively. In 2007, the US Congress established, with USD 400 million in funding, the Advanced Research Projects-Energy (ARPA-E), which was modeled on DARPA—the agency that was instrumental in funding the development of the Internet (see Bonvillian and Van Atta, 2011). ARPA-E has since its inception funded over 400 energy technology projects, which had “high-risk/high reward” profiles. The agency has been set-up to foster radical innovation in energy technology areas such as “Innovative Materials and Processes for Advanced Carbon Capture Technologies,” “Electrofuels,” “Batteries for Electrical Energy Storage,” “Agile Delivery of Electrical Power Technology,” “Grid Scale Rampable Intermittent Dispatchable Storage,” and “Building Energy Efficiency Through Innovative Thermodevices.” An additional USD 2.1 billion was invested by the US government through tax credits. Federal subsidies for renewable energy increased from USD 5.1 billion to USD 14.7 billion between 2007 and 2010. The USD 787 billion federal stimulus package, which was enacted in 2009 in response to the global financial crisis, included USD 79 billion for renewable and clean energy. Furthermore, USD 2.3 billion in tax credits were awarded to renewable and clean energy companies.

These massive government subsidies, in turn, fueled VC investments. Whereas clean-tech investment amounted to a few hundred million dollars in 2005, in the following year, VCs invested USD 1.75 Billion in clean-tech startups. Although, clean-tech does not resemble the software investments that have been dominating VCs portfolios, investors compared solar, biofuels, and batteries to the capital-intensive semiconductor and biotech industries. As Doerr remarked:

“Internet-sized markets are in the billions of dollars; the energy markets are in the trillions” (Doerr, 2007). Kleiner Perkins, for example, estimated that the total annual information technology market was USD 1 trillion a year, while that for energy was USD 6 trillion (see Yergin, 2011). In 2009, at the peak of the bubble, VCs invested USD 4.9 billion into 356 alternative energy deals. Whereas VC investments in solar increased from USD 32 million in 2004 to USD 1.85 billion in 2008, venture investments in batteries increased thirty-fold during this period (see Eilperin, 2012; Hargadon and Kenney, 2012; Gaddy et al., 2016).

VC investments in clean-tech assumed increasing fossil fuel prices, in particular, of natural gas. However, a confluence of factors, such as the financial crisis of 2008 that made it more challenging for VCs to raise capital, the decline in natural gas and oil prices, a drop in the prices of processed silicon, technological advances in natural gas extraction from shale, and overproduction in solar panel manufacturing, triggered the bust of the clean-tech bubble in 2008. One of the most dominant factors, which accelerated the bursting of the clean-energy bubble, was the decline in natural gas and oil prices, which we will analyze in more detail below. Technical advances in natural gas extraction, such as hydraulic fracking, resulted in a collapse of natural gas prices. While natural gas peaked at USD 13 per million Btu in 2008, it crashed to around USD 3 the following year. Another factor that contributed to the bust of the clean-tech bubble was the overproduction of processed silicon that resulted in a substantial decrease in silicon prices. This price-decrease, in turn, lowered costs and barriers-to-entry for other solar panel manufacturers. Many next-generation renewable energy startups—such as Solyndra and the expensive CIGS-coated cylinders they used for their novel, cylindrical solar cells—could no longer compete. Furthermore, driven by government investments and credits, China became in 2007 the largest producer of solar panels (see Lacey, 2011; Fehrenbacher, 2015). Many clean-tech startups, such as Solyndra, Abound Solar, or Evergreen Solar, later blamed low-cost Chinese competition and pricing for their failure. The increased supply of cheap Chinese solar panels also affected the adoption of wind turbines, which made them a less cost-effective clean energy technology (see Yergin, 2011).

As already mentioned, in 2008, the clean-tech bubble went bust. The Renewable Energy Industrial Index (RENIXX), which tracks the largest companies by market capitalization in the renewable energy sector, crashed in a cumulative loss of 64%. Many prominent clean-tech startups, such as Evergreen Solar, A123, or Solyndra—which required massive up-front investments to develop new hardware and scale up manufacturing—filed for bankruptcy. Solyndra, for example—which was the first clean-tech company to receive a federal loan guarantee, but defaulted on USD 527 million in federal loans—raised USD 1 billion in VC capital. Solyndra’s bankruptcy represents one of the largest losses in the history of venture capital. As a consequence of the bursting of the bubble, VC investments decreased from USD 4.1 billion in 2008 to USD 2.5 billion in 2009.

As investors started to realize, clean or renewable technologies represent a “complex, established legacy sector” (Bonvillian and Weiss, 2009). In other words, clean-tech did not satisfy the conditions for massive venture capital financing to successfully accelerate radical technological breakthroughs. These conditions consist of (i) rapidly growing markets, (ii) scalable technology and startups, and (iii) large returns and rapid exits (that is, private and public market liquidity) (see Hargadon and Kenney, 2012). VCs assumed that clean-tech constitutes a rapidly growing market, as illustrated by the statement of Doerr that clean-tech represents “the biggest economic opportunity of the 21st century.” In reality, the energy sector is characterized by high-barriers-to-entry and high-capital costs, which constraints the rapid growth and diffusion of emerging clean technologies. Federal subsidies and investments in clean energy created artificial market growth and reduced investor risk-aversion. Furthermore, technology and startup scalability is difficult due to the need for long time-horizons and for high capital intensity of “hardware”-dominated clean-tech projects. Thus, successful software-applications to the clean energy sector provide exceptions, such as Nest, which was acquired by Google in 2014 for USD 3.2 billion. The larger time-scales that clean-tech requires make it difficult for investors to achieve rapid exits. In essence, high scaling costs and the challenge for clean-tech startups to generate outsized growth in returns and market share—which VC investors need for growing their funds—make the clean energy sector difficult to invest in for VCs (see Gaddy et al., 2016). ARPA-E’s innovation-policy model played also a critical role in the clean-tech bubble. In contrast to DARPA, ARPA-E’s policy resulted in the concentrated funding of selected high-risk/high reward startups. It specifically subsidized specific technologies, which, in turn, resulted in reduced competition between emerging technologies (see Bonvillian and Weiss, 2009). In other words, ARPA-E’s policies created selected and over-funded potential winners, which raised barriers-to-entry for other clean-tech startups. If the best venture capitalist failed to pick winners in clean-tech, it may not be surprising that the government failed as well.

Before examining the social dynamics underlying the clean-tech bubble through the lens of the Social Bubble Hypothesis, we will in the next section provide a quantitative analysis to size up the clean-tech bubble.

3. Quantitative Analysis of the Clean-Tech Bubble

3.1 Super-exponential Growth and Log-Periodic Power Law Singularity as Diagnostic of Bubbles

A statistical signature of financial bubbles is the existence of “acceleration” (Ardila-Alvarez et al., 2021), i.e., increasing momentum, leading to super-exponential price growth (Sornette and Cauwels, 2015). The pattern of faster-than-exponential (power law hyperbolic) price growth, which is fueled by positive feedbacks on the growth rate of an asset’s price by the price, return, and other financial and economic variables, makes bubbles detectable, in particular through log-periodic power law singularity (LPPLS) patterns (Johansen et al., 1999; Johansen and Sornette, 2010; Sornette, 2003; Sornette and Johansen, 2001; Sornette and Zhou, 2006; Jiang et al., 2010). This signature of positive feedback loops, which are characteristic for bubble regimes, can be quantitatively identified in a time series by a faster-than-exponential power law component, and by the existence of an increasing amplitude of low frequency volatility, which occurs either in isolation or simultaneously with varying relative importance. Mathematically, such bubble regimes of unsustainable growth can be identified by power law finite-time singular growth that is decorated by oscillations in the logarithm of time. The mathematical representation is obtained as the expansion of the Log-Periodic Power-Law Singularity (LPPLS) model for the expectation of the log-price:

$$E[\ln p(t)] = A + B|t_c - t|^m + C|t_c - t|^m \cos(\omega \ln|t_c - t| + \phi), \tag{1}$$

where P(t) is the price of the asset, t is time, t_c is the most probable time at which the bubble ends, *E*[.] represents the expectation operator and A, B, C, f, m, and ω are parameters defining the nonlinear equation. The exponent 0<m<1 quantifies the strength of the super-exponential acceleration. The log-periodic angular frequency ω controls the discrete scaling structure of the accelerating bursts of volatility accompanying the development of the bubble (Sornette, 1998).

3.2 Quantifying the Clean-Tech Bubble: A LPPLS Analysis

Applying the LPPLS model (1) to the Renewable Energy Industrial Index (RENIXX), we can detect a regime of super-exponential growth in the RENIXX from August 2004 to December 2007 (see Figure 1).

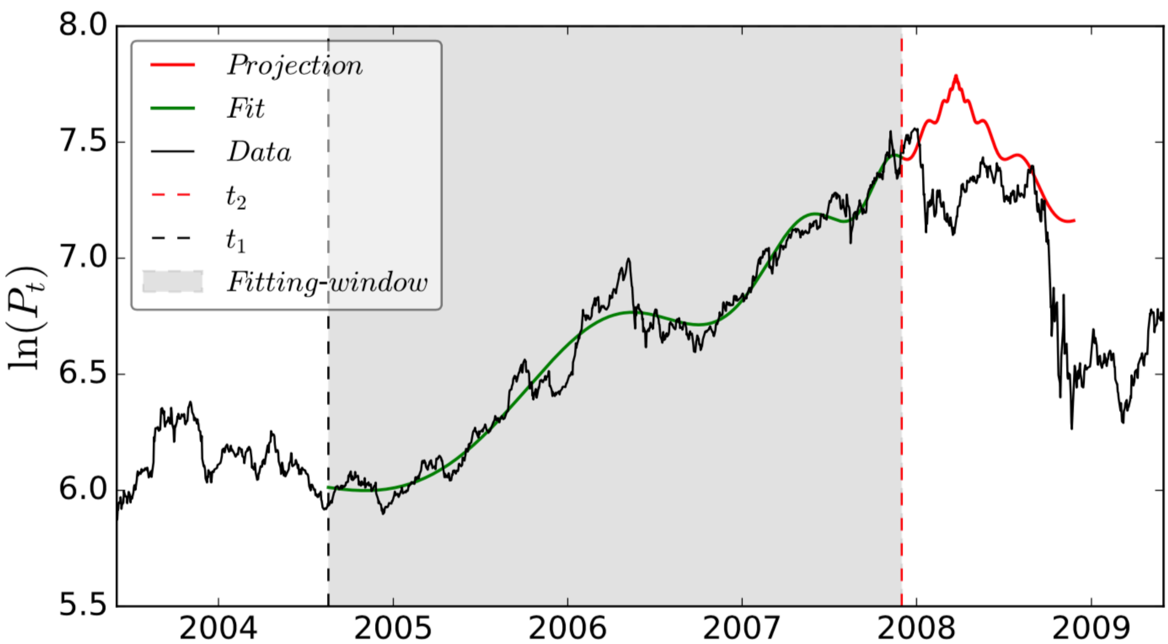


Figure 1: RENIXX daily prices in logarithmic scale from 2003 to 2009. The oscillating continuous green line corresponds to the calibration of equation (1) (with the expansion of the cosine as C cos[w ln(t_c-t)+ f] = C_1 cos[w ln(t_c-t)] cos[f] – C_2 sin[w ln(t_c-t)] sin[f], where C_1=C cos[f] and C_2=-C sin[f], as proposed by Filimonov and Sornette (2013)) to the RENIXX with the parameters: t_c = 113.5, A = 2.054, B = –0.002, C1 = 0.0003, C2 = –0.0001, m = 0.67, ω = 7.56. The fitting window is bracketed by the two vertical dashed lines from August 2004 to December 2007, and lasts approximately 1200 days. The overall upward curvature in this log (price) versus time means the price accelerated faster than exponential, matching the LPPLS definition of a bubble (see text).

The start time t_1 of the interval from August 2004 to December 2007, over which the calibration of equation (1) is performed, is selected using the method of Demos and Sornette (2019). The end time t_2 of this interval is selected between the two last largest peak of the RENIXX index. As can be seen in figure 1, the RENIXX index exhibits a regime of accelerating price corrections and rebounds—which are a defining characteristic of bubbles—that are driven by self-reinforcing feedback loops of herding behavior (Sornette, 2017). The price trajectory of the RENIXX becomes unsustainable and crashes between 2008 and 2009. The best LPPLS fit shown in figure 1 is extrapolated beyond t_2, showing a correct timing of the transition from bubble to drawdown. It is important to recall that the LPPLS model describes the end of a bubble as a regime shift, which can often extend over several weeks or months.

Within the theory of financial bubbles and their subsequent regime change, which is used here, the specific factor triggering the drawdown is only a proximate cause, and not the fundamental one. The crash has occurred because the system itself has entered an unstable phase, and any disturbance may have caused it. In other words, the systematic instability of the system itself, embodied in the super-exponential LPPLS pattern, can be identified as the underlying cause of a market crash. While factors, such as the fall in natural gas prices, the emergence of fracking, the 2008 financial crisis, or the Chinese solar industry, have all contributed, it was the unsustainability of the price growth itself that was the ultimate cause of the bursting of the clean-tech bubble. Now, as a speculative bubble is often influenced by developments occurring concomitantly in other markets, we now quickly test in next sections how the global stock market and, in particular, the price of oil has affected the clean-tech bubble.

3.3 The Clean-Tech Bubble and the Stock Market

Expanding on figure 1, figure 2 exemplifies to what extent the RENIXX went through a bubble from 2005 to 2008. The index increased in 2005 by more than 62%, in 2006 by 45%, and, at the peak year of the bubble in 2007, by more than 107%, passing from 395 points at the beginning of 2005 to 1918 points at the end of 2007. In 2008, the index crashed by 64%, stabilized in 2009 when it gained 7%, and then decreased in 2010 by another 29%. In the following year, it further lost 54% and, in 2012 it lost another 30% to attain its lowest level of 145 points in November 2012, less than half its value in 2005.

Figure 2: Renewable Energy Industrial Index (RENIXX) (daily close prices) as a function of time from January 2003 to December 2016.

Figure 3: RENIXX-in-S&P500(t)=RENIXX(t)/S&P500(t) daily close prices (in logarithmic scale) as a function of time from January 2003 to December 2016.

During 2007, the RENIXX almost continuously increased. It increased by 107.3%, reaching a high of 1918 points on December 28, 2007. However, this remarkable ascent was interrupted abruptly when the RENIXX crashed for the first time reaching a low of 1221 in early 2008, which represented a decrease of 35% since the end of the previous year. In 2008, the index started to

rise again by 35% to reach a second peak in June 2008 of 1652. This peak was followed by a drawdown of 58% in the second half of 2008. As this crash of the RENIXX overlaps with the time period of the bursting of the real estate bubble in the US and its stock market drawdown, associated with the 2008 Great Financial Crisis, the question arises whether the clean-tech bubble represented an intrinsic bubble in clean energy, or whether it was simply driven by the global market through a kind of contagion effect.

To disentangle the dynamics of the RENIXX index from that of the stock market, we can express the RENIXX "in currency units" of the Standard & Poor's 500 (S&P500), a stock market index based on the market capitalizations of the 500 largest companies listed on the NYSE or NASDAQ, which is used as a global market indicator. In other words, we take the S&P500 index as the numeraire to express the value of the RENIXX index in relative terms as

$$\mathrm{RENIXX\text{-}in\text{-}S\&P500(t)} = \frac{\mathrm{RENIXX(t)}}{\mathrm{S\&P500(t)}} \tag{2}$$

Figure 3 shows the ratio (2) in a logarithmic scale as a function of time. One can observe that this ratio increased by a factor of 4.3 (from 0.3 in December 2004 to 1.3 in January 2008) during the rise of the green-tech bubble. With the crash of the bubble, the ratio went down to 0.1 in November 2012, which represents a 12-fold decrease from the height of the bubble. Figure 3 demonstrates unambiguously that RENIXX massively over-performed the global US stock market for more than three years before underperforming it catastrophically for four years. In the aftermath of the Great Financial Crisis, the RENIXX index did not benefit from the quantitative easing programs that were launched in the end of 2008 and developed for several years thereafter. This evolution of the "RENIXX in S&P500" confirms that the RENIXX index went through an intrinsic bubble of its own, with a much larger amplitude than that of the US stock market at large. While we cannot exclude some contagion effect, the green bubble amplified extensively the exuberance of the pre-Great Financial Crisis regime, with the RENIXX index beating the S&P500 index by an average of more than 62% per year from December 2004 to January 2008.

3.4 The Oil Prize and the Clean-Tech Bubble

As clean-tech is considered to be a substitute for oil, and more generally for carbon-based energy sources, we now analyze in more detail how the oil price affected the clean-tech bubble. Before and during the clean-tech bubble, the oil price increased substantially before crashing (see Sornette et al. (2009) for a detailed LPPLS study of the oil bubble). From 2004 to 2006, the oil price increased to exceed USD 75 a barrel in mid-2006 and then dropped back to USD 60 a barrel in early 2007. The price then increased to USD 92 a barrel in October 2007 and continued to rise to the record high of USD 147.02 a barrel on July 11, 2008 before dropping very quickly thereafter. By October 2008—only three months after the historical peak—the oil price had fallen below USD 70 a barrel. This decline in oil price has been attributed to slowing demand due to the global financial crisis, but Sornette et al. (2009) demonstrated also the presence of a significant speculative bubble component, which explained the abrupt change of regime and crash.

Figure 4: The oil price per barrel in blue, on the left scale, and Renewable Energy Industrial Index (RENIXX) in red, on the right scale, (both daily close prices) as a function of time from January 2003 to December 2016.

Figure 4 shows that the RENIXX index went through three main peaks before crashing: one peak, which occurred in December 2007, and two peaks in close succession within a month, one of them slightly anticipating in June the oil price peak on July 11, 2008, and the second one a few weeks later. The first peak in December 2007 reflects the intrinsic endogenous dynamics of the clean-tech bubble, and its timing is accurately quantified by the LPPLS analysis, as presented in Figure 1 above. Within the LPPLS framework, this peak and its subsequent crash reflect the psychology of investors who realized that, given that renewable and clean energy is more complex and costly to develop than expected, the super-exponential exuberant pricing was not justified, nor sustainable. However, as the oil price was still accelerating, and analysts were extrapolating that the oil price could reach USD 200 or more within the next two years (see Menon, 2008), investors were enticed to re-adjust their valuations of clean-tech upward, leading to a short rebound of the RENIXX index after its first crash of 35%. This rebound paralleled for a while the growth of the oil price, however with a first drop after the second peak before the oil price peak, followed by a rebound (third peak). This shows the strong susceptibility of the RENIXX index over this time period, as revealed by the LPPLS analysis that identified the change of regime as the expected development following its bubble phase. Figure 4 shows an apparent synchronization of oil and RENIXX in their big crashes (see Figure 5 for a more detailed view).

The relation between the two time series during the crash following the second peak is very informative about the influence of the oil price on the RENIXX. The period of the crash is magnified in Figure 5, which suggests that both time series are closely related. This visual impression is confirmed formally by implementing the Johansen test for cointegration (Johansen, 1995) on the RENIXX index and the oil price. The cointegrated VAR model is given by

$$\Delta y_t = A(B' y_{t-1} + c_0) + c_1 + \epsilon(t) \tag{3}$$

where y_t is the vector composed of the two variables (RENIXX, oil price) at time t, A is the adjustment speed, B' is the cointegrating vector, c_0 and c_1 are two constants and $e(t)$ is the noise residual. On the sample of 195 daily prices from 27.08.2008 to 22.06.2009 shown in figure 5, the null hypothesis of no cointegration is rejected at the 99% confidence level (p-value=0.01). The clear cointegration of the drawdowns of the two times series supports the hypothesis that the second peak in the RENIXX was caused by the very high oil price. And the second drawdown of the RENIXX index, delayed by the strong oil price upsurge, seems to be freed to express its full power with the crash of the oil price.

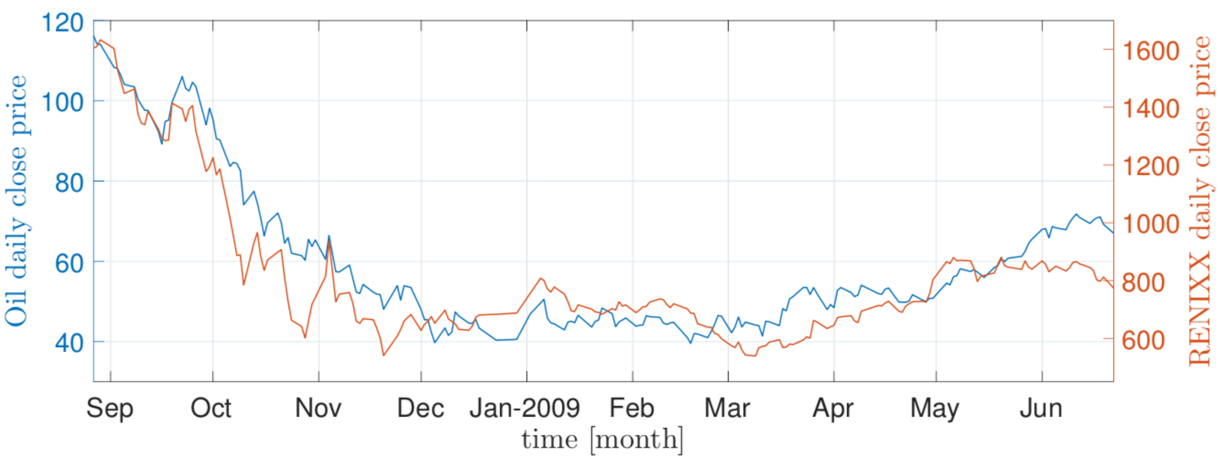


Figure 5: The oil price per barrel on the left scale in blue, and Renewable Energy Industrial Index (RENIXX) on the right scale in red, (both daily close prices) as a function of time from August 27, 2008 to June 22, 2009.

In order to further disentangle the dynamics of the RENIXX index from that of the oil price, we express the RENIXX "in currency units" of the oil price, as showed in equation (4), in the spirit of the previous equation (2),

$$\mathrm{RENIXX\text{-}in\text{-}oil(t)} = \frac{\mathrm{RENIXX(t)}}{\mathrm{oil\ price(t)}} \tag{4}$$

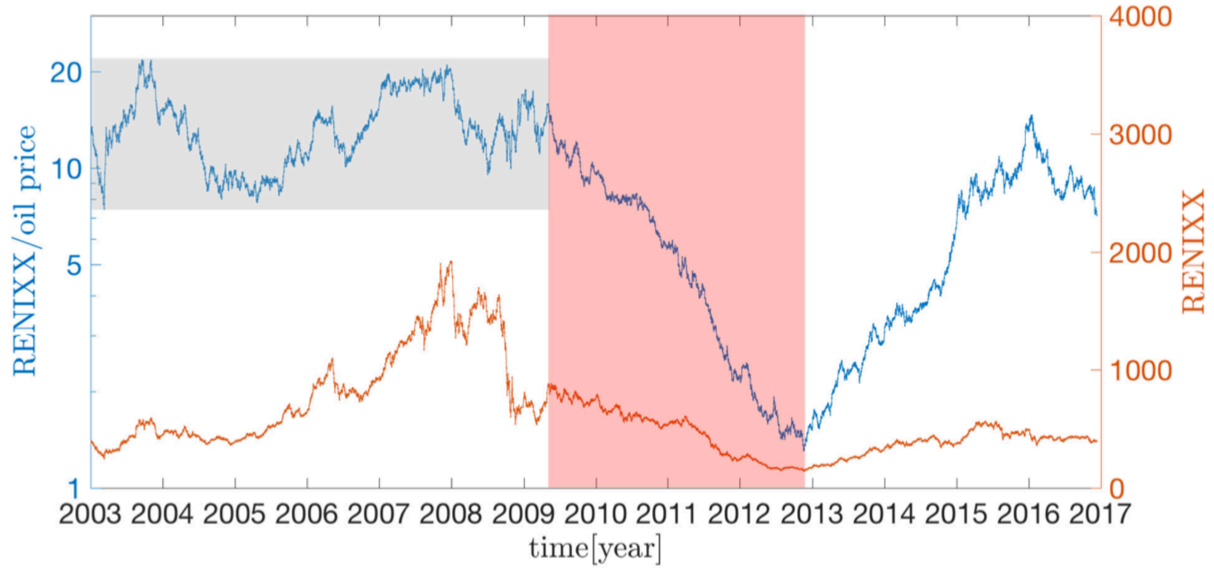


Figure 6: RENIXX-in-oil(t) defined by equation (4) (in logarithmic scale) on the left scale in blue, and Renewable Energy Industrial Index (RENIXX), on the right scale in red (both daily close prices), as a function of time from January 2003 to December 2016. The grey rectangle highlights the period from 2003 to 2009 when the RENIXX was intertwined with the oil price. The pink band highlights the period from 2009 to 2012 when the RENIXX decoupled from the oil price and the ratio RENIXX/oil price declined nearly continuously in a kind of death spiral for the RENIXX.

The grey rectangle in Figure 6 outlines the period from 2003 to 2009 during which the ratio (RENIXX/oil price) remained in a band between 8 and 20. During this period, RENNIXX/oil price increased from 8 in the early 2005 to 20 towards the end of 2007, confirming the specific bubble nature of the clean-tech sector. While oil prices had a clear influence, the RENNIXX developed a life of its own with a 2.5 fold appreciation compared with oil prices. While the market accounted somewhat for the substitutional value of alternative energy sources when the oil price surged, the clean-tech sector develops its specific dynamics. In 2009, the RENIXX decoupled from the oil price when the oil price per barrel reached over USD 100 and the RENIXX started a continuous decline (see Alsayegh, 2016). This period is indicated within the pink band in Figure 6. The ratio RENIXX/oil price went from 20 in December 2007 to 1 by the end of 2012. This massive decoupling between renewable energy and the oil price most likely resulted from investors recalibrating their expectations for clean energy in the aftermath of the burst of the clean-tech bubble and in view of the apparent broken promise of never-ending growing oil prices .

In the next section, we will now analyze the social dynamics underlying the clean-tech bubble in more detail.

4. "Salvation and Profits": The Clean-Tech Bubble as a Social Bubble

4.1 Main Elements of a Social Bubble

In the history of technology and financial markets, bubbles have often accelerated the development, diffusion, and adoption of many transformative technological innovations. As demonstrated by the railway bubble of the 1840s (see Odlyzko, 2010), the dotcom-mania of the early 2000s, or, more recently, the Bitcoin bubbles (see Huber and Sornette, 2020), bubbles can catalyze the development and adoption of emerging technologies. The exorbitant capital, which these bubbles have mobilized, funded these emerging technologies beyond what would be rationalized by standard cost-benefit analyses (Gisler and Sornette, 2009, 2011; Gisler et al., 2011). While financial bubbles, which have been at the core of technological revolutions, frequently burst, the technologies and infrastructures developed during the bubble regime find novel uses and undergo new developments in their aftermath. Bubbles occur also beyond markets in large-scale scientific or engineering projects, such as the Apollo Program (Gisler and Sornette, 2009) or the Human Genome Project (Gisler et al., 2011) demonstrate.

Financial and social bubble share a universal pattern: they result from strong social interactions between enthusiastic supporters of a technological, scientific, political, or entrepreneurial project or idea (Sornette, 2008). Exceptional enthusiasm leads to imitation and herding behavior. Fear-of-missing-out, "anticipation of regret" or aspired rewards (see Potts, 2016), as well as social signaling value, intensify the endorsement and commitment by those involved (entrepreneurs, investors, regulators, policy makers, etc.). This creates positive feedback loops, where recent price appreciations and positive developments further justify future endorsement of the hyper-optimistic narrative supporting the bubble. In such a social bubble, we can identify the following phases:

- Initial scientific or technological breakthrough, specific idea, invention, or project;
- Public and private investments, which result in increasing price or valuation growth in corresponding firms, sector, or market; this growth attracts other less sophisticated and international investors that further fuel price increases;

- Proliferation of ventures of all kinds (which often leads to the dismissal of best-practice models, such as standard risk-benefit analyses, which, in turn, reduces collective risk-aversion), leading to unsustainable accelerated price increase;
- Disappointing outcomes, reassessment and hard confrontation with reality, abrupt project or program termination.

If we now map these generic features—which we have extracted from previous case studies—onto the clean-tech bubble, it becomes evident that this bubble shared many of the ingredients of a social bubble. We will now explore in more detail the social dynamics that have fueled the clean-tech bubble.

4.2 The Narrative Economics of the Clean-Tech Bubble

To develop a deeper qualitative understanding of the clean-tech bubble, we now need to turn to the narrative economics driving the bubble. Historically, at the core of each bubble, we can identify a narrative that resulted in speculative contagions and bubble dynamics. While narratives are key factors in the emergence of bubbles, a comprehensive analysis of bubble narratives has been lacking. Only recently, economist Robert Shiller advanced what he calls narrative economics, which is the study of the "spread and dynamics of popular narratives […] to understand economics fluctuations" (see Shiller, 2017; 2020).

The clean-tech bubble is one of the purest examples of a bubble that has been driven by a narrative. The core narrative of the clean-tech bubble is encapsulated in "salvation and profits," the title of the talk venture investor John Doerr gave on clean-tech. While clean-tech represented an opportunity to be exploited for profit—an opportunity "bigger than the Internet"—it also had a quasi-religious dimensions to it: investing in clean-tech could promise "salvation." As Shiller notes, a narrative that has emotional resonance has a potential for contagion. Doerr, for example, started his talk with an anecdote that, during the talk, moved him to tears. An interaction with his daughter on how climate change will impact her generation made it clear for him that solving the problem of climate is a moral imperative: "We cannot afford to underestimate this problem. If we face irreversible and catastrophic consequences, we must act, and we must act decisively" (Doerr, 2007). Kleiner Perkins–John Doerr's VC firm that famously pivoted the fund away from investing in internet startups to clean tech–also exploited climate change and the impending crisis to market their USD 500 million clean tech growth fund as a vehicle "intended to help speed mass-market adoption of solutions to the world's climate crisis." The clean-tech narrative invariably involves a reference to "crisis" and "catastrophe." The impending climate crisis creates an immediate personal and emotion-lading connection to the narrative that fuels the contagion of the narrative at the core of the speculative bubble (see Shiller, 2017).

Similarly, Elon Musk, a co-founder of electric vehicle-maker Tesla[2], constantly mobilizes references to climate change and an impending catastrophe when he refers to the mission of his company: "Why does Tesla exist? Why are we making electric cars? Why does it matter? It's because it's very important to accelerate the transition to sustainable transport. […] This is really important for the future of the world" (Musk, 2016). In a blog post on Tesla's website, Musk writes, activating the emotional connection to the narrative: "We must achieve a sustainable energy economy or we will run out of fossil fuels to burn and civilization will collapse" (Musk, 2016). For Musk, Tesla is a key-component in this acceleration of the "transition to sustainable energy," which he envisions as a vertically-integrated clean-energy company that produces not only electric cars but also batteries and solar roofs (Musk was also involved in launching SolarCity, which develops and sells solar panels and solar roof tiles and which later became a subsidiary of Tesla). Tesla is, therefore, according to Musk "very important for the future of the world […] It's very important for all life on Earth. This supersedes political parties, race, creed, religion, it doesn't matter. If we do not solve the environment, we're all damned" (Musk, 2018). We can identify in Musk's rhetoric "salvation" and "damnation" as important imagery, which also Doerr employed and that is inexorably linked with narrative driving the clean-tech bubble, in particular, and clean-tech in general.[3]

Historically, energy and narratives have been deeply intertwined. As the example of nuclear energy illustrates, narratives about technologies, its promises and perils, fundamentally shape their trajectories. Whereas radiation was, in the decades that followed the discovery of radium in 1896, perceived as a form of "alchemy" and "transmutation" that could lead to utopian civilizational renewal, nuclear energy and radiation became, catalyzed by Hiroshima and Chernobyl, irreversibly linked with the dystopian imagery of "contamination," "mutation," and "destruction" (see Weart, 2012). The rise of nuclear fear demonstrates how powerful collectively-shared narratives and imagery is for technological adoption and diffusion. The narratives of alternative "clean" or "renewable" technologies of energy, which followed the anti-nuclear narratives of "contamination" and "doom," also powered the clean-tech bubble. "Clean" and "renewable" give rise to a spiritual, quasi-religious, or mystical imagery of purification, healing,

---

[2] While founded in July 2003 by Martin Eberhard and Marc Tarpenning as Tesla Motors, a lawsuit settlement agreed to by Eberhard and Tesla in September 2009 allows all five—Eberhard, Tarpenning, Wright, Musk and Straubel—to call themselves co-founders [Wikipedia: https://en.wikipedia.org/wiki/Tesla,_Inc., accessed 9 May 2021]

[3] As a result of Musk's mastery of investment narratives, which he developed during the clean-tech bubble, Tesla itself became a "meme-stock," that is, a stock that is primarily driven not by fundamentals but by narratives. Tesla, it could be argued, is part of a larger trend of the *memification* of markets. Tesla's market capitalization, which now exceeds the combined market cap of the world's largest car makers, the Gamestock short squeeze that was driven by WallStreetBets, a forum on the social media platform Reddit, or the rise of Dogecoin, a meme-based cryptocurrency, are recent examples of the increasing importance of memes as investable narratives in markets. Musk himself recently even tweeted that "I am become meme": https://twitter.com/elonmusk/status/1357269755112148993).

and renewal, which contrasts with the deeply entrenched and collectively-shared imagery of crisis and pollution that nuclear energy and fossil fuels invoke.

A bubble continually self-reinforces the narratives and imagery that are driving it. In the case of the clean-tech bubble, “climate change” or “clean” or “renewable” energy have become memes themselves. In the context of speculative financial bubbles, memes—understood as a self-replicating and mutating symbolic units of cultural transmission and imitation (see Dawkins, 2006)—can help investors and traders process information and navigate markets. While it can result in bubbles, imitation in markets can be rational for individual investors if there is an informational overload or lack of information. Rather than acting on private information or signals, investors imitate the behavior of other investors, which can lead to so-called informational cascades or mimetic contagions (see Sornette, 2003; Orléan, 1986). As they compress information and evoke emotions, memes, which result from a process of memetic selection, can thus encode valuable information for investors and speculators. However, the self-organizing dynamics of memes and narratives of bubbles can lead to spectacular busts and crashes.

In the following section, we will examine why the clean-tech bubble went bust. We will then address the question whether the clean-tech bubble resulted—similar to preceding innovation-accelerating bubbles—in novel technological breakthroughs and a large-scale adoption of clean technologies.

4.3 The Clean-Tech Bubble As A Social Bubble

At the core of previous social bubbles was often a specific technological breakthrough. In contrast, the speculative enthusiasm to invest in clean technologies was largely driven by the hyped threat of climate change, which was reasoned to be the harbinger of fundamental societal shifts. As mentioned above, wanting to catalyze the perceived needed energy shift, VCs and government agencies, such as ARPA-E, invested in a cluster of emerging clean technologies, such as solar, biofuels, or batteries.

After the collapse of the dotcom-bubble in the early 2000s, venture capitalists started to identify what Schumpeter termed “new economic spaces” (see Schumpeter, 1939). For these investors, clean or renewable energy represented such a new economic space in its early stages with new and potentially high-growth industries. As Doerr stated “energy markets are in the trillions of dollars.” It is in the early 2000s that Silicon Valley VCs first started to invest in clean technologies. Vinod Khosla, for example, started to invest heavily with his VC firm Khosla Ventures in biofuels and other renewables. In 2004, Elon Musk led the Series A round of investments in Tesla, which was founded in the previous year, and joined its board of directors. By 2008, Kleiner Perkins had allocated more than USD 300 million to clean-tech and launched a USD 500 million growth clean-tech fund. Between 2004 and 2009, the firm had invested USD 630 million across 54 clean-tech companies, and 12 of its 22 partners spent some or all of their time on so-called green investments. These private and public clean-tech investments triggered a positive feedback loop, which attracted new capital, entrepreneurs, and startups. The IPOs in

2005 of, for example, Q-Cells AG, SunPower Co., and Suntech Power Holdings Co. Ltd., further reinforced the interest of investors. Fueled by federal subsidies and tax credits, new clean-tech ventures proliferated. As John Doerr stated, clean-tech was not only “the largest economic opportunity of the 21” but also a “moral imperative.” Coupled with the herding and imitation behavior of Silicon Valley venture capitals, in particular resulting from the so-called “fear-of-missing-out,” government funding of the emerging clean-tech sector reduced risk-aversion and distorted market signals. While it seemed that clean or renewable energy represented a high-growth sector, the market was mainly growing because of a massive inflow of venture capital and government subsidies. In 2008, the clean-tech bubble burst, as described above, because of a confluence of factors, such as falling natural gas prices, the financial crisis, and the growth and dominance of China’s solar industry (see Gaddy et al., 2017). The VC-model—which relies on scalability, rapidly growing markets, and exit opportunities—failed in hardware-dominated clean technologies because of the capital-intensity and the long time horizons required for commercializing and scaling these emerging technologies. Clean-tech also provided limited exit opportunities: large industrial corporations did not acquire clean-tech startups as it is the case in the IT or biotech industries. However, VCs continued to invest in clean-tech after the burst of the bubble, but—in order to avoid the “Valley of Death” and reduce technological risks—mainly at later stages and in clean-tech startups that are less capital intensive and have defined go-to-market strategies.

The enthusiasm for clean-tech turned out to be inflated. Investors and entrepreneurs in the early phase of the bubble developed extraordinarily over-optimistic expectations about the total addressable market and the future large-scale adoption of clean or renewable technologies. One could argue that the bubble in clean technologies was, at a deeper level, driven by what venture capitalist Peter Thiel has termed “indefinite optimism.” In other words, what was fueling the clean-tech bubble was an indefinite vision of the future. Whereas previous bubbles were premised on a specific technology or a set of interrelated technologies, the clean-tech bubble was driven by investments in various different and, often, unrelated technologies, such as solar panels, batteries, or biofuels. While it was in most cases unclear how these novel technologies could be produced, commercialized, scaled, and diffused, very large amounts of capital was nevertheless flowing into clean technologies and startups. Similarly, investors and entrepreneurs had unrealistic expectations about the size of the total addressable market—they were indefinitely optimistic about the transformative potential of clean-tech technologies and size of the market they can capture (see Thiel and Masters, 2012). Driven by hype, imitation behavior among investors and entrepreneurs, and massive government subsidies, which reduced collective risk aversion and distorted market signals and incentives, investors and entrepreneurs took inordinate risks that would not otherwise be justified by standard due diligence processes, cost-benefit and portfolio analyses. Clean-tech also turned out be more expensive than traditional energy technologies. Whereas the price of solar panels per watt declined by 75 percent between 2009 and 2017 and the price of wind turbines per watt declined by 50 percent, electricity prices increased between 21% and 51% (Fu et al., 2017). Solar photovoltaic, for example, is 10 times more capital intensive than nuclear energy. Studies that quantified the efficiency of an energy production technology with the ratio of Energy Return over Energy Invested (ERoEI) show that most clean-technologies are less efficient than traditional energy technologies (see Ferroni and Hopkirk, 2016 and, for a complementary assessment, Brockway et al., 2019). As a consequence, investors' unrealistic expectations about the future of clean technology crashed and VC investments in the sector decreased from USD 4.1 billion in 2008 to USD 2.5 billion in 2009.

The clean-tech bubble was clearly a social bubble: the narrative of a "moral imperative" to combat climate change and achieve "salvation," the ballooning venture capital investments, and the massive government subsidies weaved a network of self-reinforcing spirals that led to over-optimistic expectations, excessive enthusiasm, and over-investments. The question now arises whether the clean-tech bubble was—as it has been historically the case for a number (but not all) bubbles—accelerated the development, deployment, and diffusion of clean technologies. In other words, did viable commercial and industrial infrastructures and products emerge after the bust of the bubble?

5. Seeing Green?

As mentioned above, although speculative market bubbles can be excessive and have resulted in many failed investments in the short or medium term, many bubbles significantly accelerated the development and diffusion of emerging technologies and infrastructures in the longer term. Although the clean-tech bubble went bust, we can identify some factors that indicate that the bubble did indeed catalyze technological progress in clean and renewable energy technologies. The global capacity of solar photovoltaic, for example, increased from 6 installed gigawatts in 2006 to 303 installed gigawatts in 2016 globally, which represents an increase by a factor of 50 over 10 years. Furthermore, the additional installed capacity has been—except for 2011—increasing every year since 2006. Global wind power capacity progressed by a factor of 6.5 from 2006 to 2016, increasing from 74 to 487 installed gigawatts. By the end of 2016, more than 90 countries had been actively involved in commercial activities surrounding wind power, and at least 24 countries met 5% or more of their annual electricity demand with it. Moreover, offshore wind power had a global installed capacity of 14 gigawatts in 2016 (see Sawin et al., 2017). And the global capacity from combined solar PV and wind power increased by a factor of 10 between 2006 and 2016.

Over the last decade, costs for renewable and clean energy technologies also started to significantly decrease. In 2009, the levelized cost of solar photovoltaic electricity was $359 per megawatt-hour—more than four times as expensive as electricity from a natural gas plant (see also Victoria et al., 2021). By 2019, solar PV had fallen in price to USD 40 per megawatt-hour, 28% cheaper than gas, which represents a 89% decline over 10 years. The price of electricity from utility-scale solar projects has dropped by a factor of at least 5. By decreasing by 80%, the building of new solar capacities has become cost-competitive with building new coal or gas power plants across most of the world (see Naam, 2020). This price decrease is decades ahead of what forecasters predicted. In 2020, the price of solar reached prices that the International

Energy Agency, the world's foremost energy authority, did not predict until 2035, that their 2014 solar roadmap did not forecast until after 2050, and that the IEA's 2010 forecast did not expect solar to ever achieve (see Naam, 2020). In other words, similar to the exponential increase in computing power that is referred to as Moore's Law, solar technology experienced, in the decade that followed the clean-tech bubble bust, an exponential decrease in costs. This exponentially decreasing cost curve—or increasing learning curve—is captured by Wright's Law, which models an exponential decline in the cost of technologies as a function of the cumulative scale of production.[4] In the case of solar, it is well-established that the price of solar modules per watt of power drops by around 25% for every doubling of cumulative manufacturing. Even when accounting for the costs of overall solar systems, which include mounting systems, tracking systems, DC-to-AC inverters, cabling, etc., the cost of the overall solar technology follows an exponentially decreasing curve (see Naam, 2020). Lithium ion batteries, which are an integral part of electric vehicles, also experienced a similar drop in prices (see Lee, 2020).

Another startup and industry that emerged successfully from the bust of the clean-tech bubbles are Tesla and electric vehicles, respectively. By starting with a very small submarket—the market for electrically powered sports cars—then expanding into other submarkets with cheaper models and, ultimately, into a vertically-integrated clean energy business, Tesla validated clean-tech and climate change as viable investment categories.[5] Furthermore, Elon Musk demonstrated that Silicon Valley is capable of building large-scale and capital-intensive businesses, such as car companies. In other words, after the bursting of the bubble, Musk and Tesla became symbols for clean-tech and a model that could be imitated by other entrepreneurs and investors. And indeed, funding for battery and electric vehicles startup has started to accelerate in this decade—to the extent that there are mounting concerns about another clean-tech bubble in the making. However, this time, the dominant narrative is less about "clean tech" but more about "sustainability," "climate tech," and "ESG."[6] Whereas investors invested around USD 16 billion in 2015, investors invested more than USD 36 billion into climate-related technology in 2019. For comparison, in the aftermath of the clean-tech bubble, clean or climate tech attracted USD 418 million from VCs, which represents three times the growth rate of venture investments into artificial intelligence. Many new private equity and VC funds, foundations, and family offices have been launched over the past few years that invest only in clean technologies. Breakthrough Energy Ventures, founded in 2015, for example, is a USD 1 billion vehicle that invests only in startups with the potential to cut annual greenhouse-gas emissions by at least the equivalent of

---

[4] According to Nagy et al. (2013), this exponential cost curve decrease can be detected in at least 60 other technologies.

[5] For the factors of why Tesla succeeded and other clean-tech startups went bust in the bubble, see Thiel and Masters (2014).

[6] There are concerns that as bubble is forming in ESG investing, which stands for Environmental, Social and Corporate governance. During 2020, flows into sustainable open-end and ETF funds in the US reached USD 51.1 billion. That was a significant increase over 2019, when flows were USD 21.4 billion, and a nearly tenfold increase over 2018, when flows were USD 5.4 billion (see Morningstar, 2021).

half a gigaton of CO_2—some 1% of the world's total. Other funds provide new financing models that finance capital-intense clean-tech and climate startups so that they can escape the "valley of death."

There are a few factors that differentiate clean-tech investments in this decade from the bubble that formed in 2004. For example, lower yields force investors into long-duration assets, such as green infrastructure, which, due to ESG, have become more attractive than oil & gas. The rise of "patient" or "impact" capital also extended time-horizons for capital- and time-intense clean energy investments. In contrast to the last clean-tech bubble, which was fueled by venture capital investors and government funding, larger investors, such as private equity firms or corporations, are now investing in capital-intensive clean technologies, such as solar and wind, or carbon capture technologies. VCs, in turn, started to invest more in differentiated and lower-cost sectors, such as solar services and financing, lab-grown meat, climate-related software, or electric vehicles. Whereas there was a lack of acquisitions by large companies in 2008—which, as we have note above, reduced exit-opportunities for investors—the recent SPAC wave fueled a series of clean-tech exits and acquisitions. More than two dozens clean-tech startups, ranging from clean transportation, such as Nikola Corporation, to battery companies like Stem and Eos Energy Storage have, in 2020, gone public through SPACs, that is, special purpose acquisition companies that acquire and take company public in reverse mergers. Moreover, large tech companies, such as Google, Apple, Microsoft, or Amazon, have started to invest in transforming their energy system and infrastructure, such as data centers, telecom networks, and devices, which, in turn, has accelerated the market for clean and renewable energy.

In essence, the clean-tech bubble of the mid-2000s catalyzed a massive decrease in cost by excessively funding research and development in different clean-tech sectors, such as solar or wind. While almost all of the clean-tech startups of the last bubble failed, the clean-tech bubble, by decreasing prices and funding innovation, massively de-risked clean and renewable energy technologies. Solyndra, for example, failed because it was trying to market a cutting-edge new solar cell, which ended up being too expensive when the design costs started to decrease. Today, solar or wind are no longer risky technologies and are now even cost-competitive with legacy energy sources, such as gas or goal. This decrease in costs and the elimination of technical risks of clean tech is now catalyzing more investment opportunities, which, in turn, attracts new entrepreneurs and investors, such as Softbank, Founders Fund, Sequoia Capital, Y Combinator, and the two funds that were already investing in first clean-tech boom-and-bust cycle, Kleiner Perkins and Khosla Ventures.

As we have shown above, bubbles can finance the development and diffusion of cutting-edge technologies and infrastructure. Another important feature of bubbles is that they parallelize innovation. Bubbles become accelerators for innovation if they enable self-reinforcing feedbacks loops between technologies, capital, and startups. Such a reflexive dynamic existed, for example, during the dotcom bubble in which internet startups subsidized each other. Currently, such a self-validating dynamic can be observed between software and semiconductors as every new generation of GPUs encourages more GPU-intensive applications, which, in turn, encourage semiconductor manufacturers, such as Nvidia, to produce a next generation of GPUs (see also Hobart, 2020). However, such reflexive feedback loops were largely lacking in the clean-tech bubble of 2004, which, to a large extent, was subsidized by government funding. Back then, most clean-tech startups were in a research-and-development stage and there was no market for these bleeding-edge technologies: Advanced biofuels, thin-film solar companies, and all sorts of energy storage startups of the last bubble, for example, were simply too immature and too expensive to be commercialized (see Weyant et al., 2018). However, more than a decade after the bursting of the first clean-tech bubble, such reflexive feedback loops are now starting to emerge in clean and renewable energy technologies. For example, since solar power and batteries are complements, they started to drive down each other's unit cost. Every cost decrease in solar panels consequently increases the available market for batteries, while every increase in battery manufacturing capacity increases the market for solar.[7] While it did not become evident during or immediately after the bubble, the clean-bubble seems to have significantly accelerated funding of research and development of clean technologies, which are now getting harnessed in a second wave of clean or climate investing.[8]

One of the most important properties of financial bubbles is their resemblance with self-fulfilling prophecies. For example, Tesla, which emerged from the bust of the first clean-tech bubble, is realizing its vision of "accelerating the world's transition to sustainable energy." The firm single-handedly created markets for electric vehicles and batteries. Over 250 firms are now manufacturing electric vehicles and more than 47 battery factories are currently under construction. Not only did Tesla catalyzed the founding of new electric car makers and battery startups, large legacy car makers are expected to invest up to USD 500 million into electric vehicles over the next five years. While the bursting of the clean-tech bubble in 2009 unquestionably resulted in massive losses for investors—an estimated USD 12 billion were lost (see Weyant et al., 2018)—novel technological breakthroughs, massively cost-reduced wind and solar technologies, and new industries between, for example, solar and electric-vehicle manufactures and battery producers, have been emerging from its debris. If the problem of

[7] A potential future reflexive dynamic between bubbles or technologies might be the feedback loop between Bitcoin mining and renewable energy. Bitcoin mining—which is essentially a computationally-intensive process of monetizing energy—can incentivize the development of new and cleaner energy sources, which will attract new miners that, in turn, increase the security of the Bitcoin network (see also Huber and Sornette, 2020; Square, 2021).

[8] Now, this is not to argue that the clean-tech bubble did not have any negative effects. While the need for energy alternatives was recognized by investors, the massive inflow of capital into highly complex, high-risk, and, often, experimental technology—which was driven by an emotionally-laden narrative about climate change—has diverted investment and interest away from nuclear energy, for example. Instead of funding, for example, novel system-designs in nuclear energy—which use molten salt, alternative fuels, or small modular reactors—investors in the clean-tech bubble allocated capital to often more expensive and less efficient energy alternatives (see Sornette et al., 2019; Shellenberger, 2020).

climate change and abundant energy needs to get solved, it almost seems that we might need another clean-tech bubble–one that avoids the failures of the first bubble that burst in 2009.

5. Conclusion

In the previous sections, we have synthesized the history and provided a quantitative analysis of the clean-tech bubble. We identified the possible causes that led to the bursting of the bubble in 2008, such as the 2008 financial crisis, competition from China's nascent solar industry, or falling natural gas prices, as well as the intrinsic unsustainable nature of the super-exponential stock market price growth with extraordinary expectations. We further analyzed the bubble through the lens of the Social Bubble Hypothesis, which holds that bubbles can be essential components in the process of socio-technological innovation. In particularly, we examined the narratives that were critical in fueling the bubble. In the last section of the paper, we addressed the question whether the clean-tech bubble accelerated the large-scale adoption of clean technologies and showed that adoption, for example, of wind energy or solar photovoltaic has increased in the aftermath of the bubble's burst. We argued that, while the bursting of the bubble resulted in massive losses for investors, the cost decreases that the bubble has catalyzed are now starting to make the clean or climate technologies viable alternative energy sources. We showed that the self-reinforcing feedback loops between technologies, capital, and startups are essential for innovation-accelerating and socially transformative bubbles. Given that the clean-tech bubble was largely driven by government funding, and markets for the technologies were mostly lacking, we have argued that these self-reinforcing and reflexive dynamics were absent in the first iteration of the clean tech bubble, but are now starting to emerge in a new wave of clean or climate tech investing, which is harnessing some of the technologies and cost-reductions that the first bubble has enabled.

References

Alsayegh, O.A., Ramification of Oil Prices on Renewable Energy Deployment Engineering and Technology, *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering* 10(6),1786–1790 (2016).

Aral, S., *The Hype Machine: How Social Media Disrupts Our Elections, Our Economy, and Our Health—and How We Must Adapt*, Currency (2020).

Ardila-Alvarez, D., Z. Forró and D. Sornette, The Acceleration Effect and Gamma Factor in Asset Pricing, *Physica A: Statistical Mechanics and its Applications* 569, 125367 (2021).

Battisti, R. and A. Corrado, Evaluation of Technical Improvements of Photovoltaic Systems Through Life Cycle Assessment Methodology, *Energy* 30 (7), 952-967 (2005).

Bonvillian, W.B. and R. Van Atta, ARPA-E and DARPA: Applying the DARPA Model to Energy Innovation, *The Journal of Technology Transfer* 36.5, 469 (2011).

Bonvillian, W. B., and C. Weiss, Stimulating Innovation in Energy Technology, *Issues in Science and Technology* 26.1, 51-56 (2009).

Brockway, P.E., A. Owen, L. I. Brand-Correa and L. Hardt, Estimation of Global Final-Stage Energy-Return-On-Investment for Fossil Fuels with Comparison to Renewable Energy Sources, *Nature Energy* 4, 612-621 (2019).

Chauvin, K., D. Laibson and J. Mollerstrom, Asset Bubbles and the Cost of Economic Fluctuations, *Journal of Money, Credit and Banking* 43 (Suppl. 1), 233-260 (2011).

Dawkins, R., *The Selfish Gene,* 30th Anniversary Edition (3rd ed.), Oxford University Press (2006).

Demos, G. and D. Sornette, Comparing Nested Data Sets and Objectively Determining Financial Bubbles' Inceptions, *Physica A: Statistical Mechanics and its Applications* 524, 661-675 (2019).

Doerr, J., Salvation (and Profit) in Greentech, Ted Talk, https://www.ted.com/talks/john_doerr_sees_salvation_and_profit_in_greentech (2007). Accessed: May 18, 2021

Eilperin, J., Why the Clean Tech Boom Went Bust, *Wired Magazine* (January 2012).

Farmer, R.E.A., The Stock Market Crash Really Did Cause the Great Recession, *Oxford Bulletin of Economics and Statistics* 77 (5), 617-633 (2015).

Fehrenbacher. K., China is Utterly and Totally Dominating Solar Panels, http://fortune.com/2015/06/18/china-is-utterly-and-totally-dominating-solar-panels/ (2015). Accessed: May 18, 2021

Ferroni, F. and R.J. Hopkirk, Energy Return on Energy Invested (ERoEI) for Photovoltaic Solar Systems in Regions of Moderate Insolation, *Energy Policy* 94, 336-344 (2016).

Filimonov, V. and D. Sornette, A Stable and Robust Calibration Scheme of the Log-Periodic Power Law Model, *Physica A: Statistical Mechanics and its Applications* 392 (17), 3698-3707 (2013).

Fu, R., D. Feldman, R. Margolis, M. Woodhouse and K. Ardani, US Solar Photovoltaic System Cost Benchmark: Q1 2017, *EERE Publication and Product Library* (2017).

Gaddy, B.E., V. Sivaram, T. B. Jones and L. Wayman, Venture Capital and Cleantech: The Wrong Model for Energy Innovation, *Energy Policy* 102, 385-395 (2017).

Garber, P.M., *Famous First Bubbles: The Fundamentals of Early Manias,* MIT Press (2001).

Giorgis V., Use the Hype to Innovate: Social Bubbles to Save the World from Climate Change, Master Thesis at ETH Zurich (November 2017).

Gisler, M. and D. Sornette, Bubbles Everywhere in Human Affairs, chapter in book entitled *Modern RISC-Societies. Towards a New Framework for Societal Evolution*, L. Kajfez Bogataj, K.H. Mueller, I. Svetlik, N. Tos (eds.), Wien, Edition Echoraum 137-153 (2010).

Gisler, M. and D. Sornette, Exuberant Innovations: The Apollo Program, *Society* 46, 55-68 (2009).

Gisler, M., D. Sornette and R. Woodard, Innovation as a Social Bubble: The Example of the Human Genome Project, *Research Policy* 40, 1412-1425 (2011).

Hargadon, A.B., and M. Kenney, Misguided Policy? Following Venture Capital into Clean Technology, *California Management Review* 54.2, 118-139 (2012).

Hobart, B., How Bubbles and Megaprojects Parallelize Innovation, *The Diff*, https://diff.substack.com/p/how-bubbles-and-megaprojects-parallelize (2020). Accessed: May 18, 2021

Huber, T. and D. Sornette. Boom, Bust, and Bitcoin: Bitcoin-Bubbles As Innovation Accelerators, *Swiss Finance Institute Research Paper* 20-41 (2020).

Huber, T. and D. Sornette, Can There Be a Physics of Financial Markets? Methodological Reflections on Econophysics, *European Physical Journal Special Topics* 225 (17), 3187-3210 (2016).

Janeway, W. H., *Doing Capitalism in the Innovation Economy,* Cambridge University Press (2012).

Lee, T.B., The Story of Cheaper Batteries, From Smartphones to Teslas, *Ars Technica*, https://arstechnica.com/tech-policy/2020/09/the-story-of-cheaper-batteries-from-smartphones-to-teslas/ (2020). Accessed: May 18, 2021

Morningstar, A Broken Record: Flows for U.S. Sustainable Funds Again Reach New Heights, https://www.morningstar.com/articles/1019195/a-broken-record-flows-for-us-sustainable-funds-again-reach-new-heights (2021). Accessed: May 18, 2021

Kindleberger, C., *Manias, Panics, and Crashes: A History of Financial Crises*, Springer (2017).

Musk, E., Tesla Unveils Model 3, https://www.youtube.com/watch?v=Q4VGQPk2Dl8 (2016). Accessed: May 18, 2021

Lacey, S., How China Dominates Solar Power, https://www.theguardian.com/environment/2011/sep/12/how-china-dominates-solar-power (September 12, 2011). Accessed: May, 9 2021

Musk, E., Master Plan, Part Deux, https://www.tesla.com/blog/master-plan-part-deux (2016). Accessed: May 18, 2021

Musk, E., Elon Musk: Tesla's Work 'supersedes political parties, race, creed, religion, CNBC, https://www.cnbc.com/2018/11/05/elon-musk-teslas-work-is-important-to-the-future-of-the-world.html (November 2018). Accessed: May 18, 2021

Naam, R., Solar's Future is Insanely Cheap, https://rameznaam.com/2020/05/14/solars-future-is-insanely-cheap-2020/ (2020). Accessed: May 18, 2021

Nagy B., J. D. Farmer, Q.M. Bui, and J.E. Trancik, Statistical Basis for Predicting Technological Progress, *Plos One*, https://doi.org/10.1371/journal.pone.0052669 (2013).

Nairn, A., *Engines That Move Markets: Technology Investing from Railroads to the Internet and Beyond*, John Wiley & Sons (2002).

Odlyzko, A., *Collective Hallucinations and Inefficient Markets: The British Railway Mania of the 1840s*, University of Minnesota (2010).

Orléan, A., Mimetic Contagion and Speculative Bubbles, *Theory and Decision* 27, 63-92 (1989).

Pang, A., Media Hype in Times of Crises: Nature, Characteristics and Impact on Organizations, *Asia Pacific Media Educator* 23 (2), 309-336 (2013).

Perez, C., *Technological Revolutions and Financial Capital,* Edward Elgar Publishing (2003).

Raugei, M., S. Sgouridis, D. Murphy, V. Fthenakis, R. Frischknecht, C. Breyer, P. Stolz, Energy Return on Energy Invested (ERoEI) for Photovoltaic Solar Systems in Regions of Moderate Insolation: A Comprehensive Response, *Energy Policy* 102, 377-384 (2017).

Menon, S., "Super-Spike" could Lift Oil to $200: Goldman, http://www.reuters.com/article/us-oil-goldman-idUSL0691448820080506 (May 2008). Accessed: May 18, 2021

Sawin, J.L., K. Seyboth and F. Sverrisson, *Renewables 2017 Global Status Report*, Paris: REN21 Secretariat (2017).

Schumpeter, J.A., *Business Cycles. Vol. 1,* McGraw-Hill (1939).

Shellenberger, M., *Apocalypse Never: Why Environmental Alarmism Hurts Us All*, Harper (2020).

Shiller, R., Narrative Economics, *American Economic Review* 107.4, 967-1004 (2017).

Shiller, R., *Irrational Exuberance,* Revised and Expanded Third Edition, Princeton University Press (2015).

Shiller, R., *Narrative Economics: How Stories Go Viral and Drive Major Economic Events,* Princeton University Press (2020).

Johansen, S., *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, University Press on Demand (1995).

Sornette, D. and P. Cauwels, Financial Bubbles: Mechanisms and Diagnostics, *Review of Behavioral Economics* 2 (3), 279-305 (2015).

Sornette, D., Discrete Scale Invariance and Complex Dimensions, *Physics Reports* 297 (5), 239-270 (1998). (extended version at http://xxx.lanl.gov/abs/cond-mat/9707012)

Sornette, D., Nurturing Breakthroughs; Lessons from Complexity Theory, *Journal of Economic Interaction and Coordination* 3, 165-181 (2008).

Sornette, D., R. Woodard and W.-X. Zhou, The 2006-2008 Oil Bubble: Evidence of Speculation, and Prediction, *Physica A: Statistical Mechanics and its Applications* 388, 1571-1576 (2009).

Sornette, D., W. Kröger and S. Wheatley, *New Ways and Needs for Exploiting Nuclear Energy*, Springer (2019).

Sornette, D., *Why Stock Markets Crash (Critical Events in Complex Financial Systems)*, Princeton University Press, first edition 2002, second edition (2017).

Lacey, S., How China Dominates Solar Power, https://www.theguardian.com/environment/2011/sep/12/how-china-dominates-solar-power (September 2011). Accessed: May 18, 2021

Square, Bitcoin is Key to an Abundant, Clean Energy Future, *Bitcoin Clean Energy Initiative Memorandum* (April 2021).

Thiel, P.A. and B. Masters, *Zero to One: Notes on Startups, or How to Build the Future*, Broadway Business (2014).

Potts, J., Hype as a Public Good for Innovation, Available at SSRN 2934675 (2017).

Vasterman, P., C.J. Yzermans, and A.J.E. Dirkzwager, The Role of the Media and Media Hypes in the Aftermath of Disasters, *Epidemiologic Reviews* 27 (1), 107-14 (2005).

Victoria, M., N. Haegel, I.M. Peters, R. Sinton, A. Jäger-Waldau, C.D Cañizo, C. Breyer, Solar Photovoltaics is Ready to Power a Sustainable Future, *Joule* (2021).

Weart, S.R., *The Rise of Nuclear Fear,* Harvard University Press (2012).

Weyant, J. P., E. Fu, and J. Bowerscok, *Renewed Energy: Insights for Clean Energy's Future*, Kauffman Fellow Press (2018).

Yergin, D., *The Quest: Energy, Security, and the Remaking of the Modern World*, Penguin, (2011).

Swiss Finance Institute

Swiss Finance Institute (SFI) is the national center for fundamental research, doctoral training, knowledge exchange, and continuing education in the fields of banking and finance. SFI's mission is to grow knowledge capital for the Swiss financial marketplace. Created in 2006 as a public–private partnership, SFI is a common initiative of the Swiss finance industry, leading Swiss universities, and the Swiss Confederation.

swiss:finance:institute

c/o University of Geneva, Bd. Du Pont d'Arve 42, CH-1211 Geneva 4
T +41 22 379 84 71, rps@sfi.ch, www.sfi.ch

5. Conclusion

The aim of this thesis is to contribute to a deeper understanding of the essence and dynamics of technological innovation. This section summarizes the key-findings of each paper, highlights their limitations, and identifies further areas of research.

The key contribution of the first research paper is its conceptual reconstruction of how the interaction between scientific fields can advance our scientific understanding and—as it is the case with econophysics—can even give rise to a novel field. The paper traces the rich history of cross-fertilization between physics and economics and identifies how certain models and methods, which have been imported from physics, have conceptually limited the scientific understanding of financial markets. However, given the multi-level, dynamic, and complex nature of financial markets, the paper argues—by using the example of recent approaches that have been inspired by evolutionary biology—that future advances in our scientific understanding of markets also need to incorporate scientific insights from other fields beyond physics. One of the paper's limitations is that it only provides a selective analysis of the history of cross-fertilization between economics and finance. It would be interesting to expand the historical analysis. Moreover, while the paper alludes to the differences in theory and model construction in these two fields, it could also be philosophically productive to situate the analysis in contemporary debates in the philosophy of science about the nature of scientific models and explanations. Given that models in econophysics—such as the Log-Period Power Law Singularity model of market crashes, or similar models of critical phase transitions—are used to explain universal patterns across systems that are heterogeneous at smaller scales, they provide philosophically interesting case studies of the explanatory features of models that capture the universal dynamics and features of a variety of different systems.

The second paper, which advances the Social Bubble framework, shows—on the basis of Bitcoin—how critical the dynamics of speculative bubbles are for the development, adoption, and diffusion of emerging technologies. While the existing literature on Bitcoin predominately focuses on its technical dimension, the paper demonstrates that Bitcoin is equally a social phenomenon: social dynamics, which crystallize themselves in Bitcoin's idiosyncratic culture, fundamentally shape the adoption and diffusion of the protocol and cryptocurrency. By applying and extending the Social Bubble Hypothesis, the aim of the paper is to extract valuable and generalizable insights from the history of Bitcoin, which can advance our understanding of the generic dynamics and structure of future technological revolutions. As the paper demonstrates, since its genesis in 2009, the historical evolution of Bitcoin has progressed through the phases that the Social Bubble Hypothesis identifies. Bitcoin history starts with Satoshi Nakamoto, Bitcoin's pseudonymous creator, who launched the radically novel technology on a cypherpunk mailing list and online forums. The extreme commitment and enthusiasm of its earliest adopters, then, triggered Bitcoin's ever-accelerating boom-and-bust cycles, which, gradually increased its cohorts of adopters. As the paper highlights, the extreme beliefs of these enthusiastic supporters, which evangelize Bitcoin, created a self-fulfilling prophecy that has continually attracted new developers, entrepreneurs, and speculators. A fundamental problem, which our analysis encounters, is the singular nature of Bitcoin. As the paper emphasizes, Bitcoin is a technological singularity. So, the question emerges to what extent Bitcoin's singular genesis can even be replicated. A key generalizable insight, which the paper derives from the Bitcoin case study, is that potential bubble-dynamics of emerging technologies need to be harnessed. As the case of Bitcoin demonstrates, a definite vision of the future is an essential feature around which bubbles and hype cycles can form. As our research on the Social Bubble Hypothesis has demonstrated, these hype-dynamics, in turn, can trigger the self-reinforcing feedback loops of commitment that reduce collective risk-aversion and attract new capital, believers, and adopters.

In the last paper, we extend the Social Bubble framework to the clean-tech bubble that has occurred in the last decade in green and renewable technologies. As the paper shows, the clean-tech bubble—fueled by public and private over-investment—also progressed through the stages of a prototypical social bubble, only to spectacularly burst in the final phase. In contrast to the Bitcoin bubble, however, the clean-tech bubble did not form around one specific novel technology, but rather coalesced around a cluster of inter-related and inter-dependent technologies, such as solar and batteries, wind, or bio-fuels.

While the paper concludes that the clean-tech bubble resulted in a massive de-risking of bleeding-edge technologies, a fundamental change in narrative that surrounds clean technologies, and a significant decrease in costs, one limitation of the paper is that is does not explore in more detail the negative consequences of the clean-tech bubble. A decade after the clean-tech bubble has burst, the techno-economic challenges that clean or renewable energy confront have become more evident. A problem with solar and wind, for example, is that they are too unreliable and have a low energy density. Solar and wind farms require between 400 and 750 times more land than nuclear and natural gas plants. Furthermore, because of physical aging of crystalline Si-based and thin-film-based Photovoltaic panels, the output of solar panels declines between 0.5 and one percent every year and they, as well as wind turbines, need to be replaced roughly every two decades. In contrast, nuclear plants are expected to be functional for more than 80 years. Because sunlight has a low energy density, large surfaces are needed to obtain significant power. Thus, solar farms are among the most extractive of all energy resources, requiring for instance 17 times the resources as nuclear while returning just 7-8 times the energy invested at temperate latitudes. Germany, for example, has invested USD 36 billion per year on renewables over the last five years—yet they only increased the share of electricity from solar and wind by 10 percentage points. Furthermore, the cost of renewables remains very high. Solar panels with storage deliver just 1.6 times as much energy as is invested as compared to the 50 to 75 times more energy typically delivered with nuclear.

While investors recognized the need for energy alternatives, an alternative conclusion could be that the massive inflow of capital into highly complex, high-risk, and, often, experimental technology, which occurred in the clean-tech bubble, has diverted investment and interest away from nuclear energy, which is more reliable and efficient. Instead of funding, for example, novel system-designs in nuclear energy—which use molten salt, alternative fuels, or small modular reactors—investors in the clean-tech bubble allocated capital to more expensive and less efficient energy alternatives. In other words, the clean-tech bubble had a high economic opportunity cost from the perspective of those who argue that nuclear energy represents currently the most viable substitute for fossil fuels. Rather than catalyzing nuclear innovation, a more pessimistic interpretation would be that the clean-tech bubble mostly funded clean technologies that are still limited in providing reliable sources of non-carbon-based energy. In other words, on this interpretation, the clean-tech bubble exemplifies how the misallocation of resources, which can result from speculative bubbles, can have detrimental societal effects. However, while it is unlikely that wind and solar could completely replace fossil energy sources, it is also not likely that we can rely exclusively on nuclear energy. Rather, it is more likely that a broad portfolio approach, including nuclear energy and full-cycle gas in addition to wind and solar, and other sources appropriate to local conditions (like geothermal), will constitute the energy solution of the future. However, given that the clean-tech bubble fundamentally shifted the narrative around clean-technologies and financed and de-risked emerging clean and renewable technologies, the paper's conclusion that the clean-tech bubble was societally net-positive remains valid.

An interesting area of future research, which has not been developed further in this thesis, is the question to what extent bubbles can be engineered. This question has practical ramifications for entrepreneurs as well as policy makers. As the clean-tech case study shows, policy-decisions related to the funding of emerging technologies—which, in the case of the clean-tech bubble, resulted in the concentrated funding of selected high-risk/high reward startups—can distort market signals. Now, given that it provides a powerful conceptual framework that identifies the key-components of bubble-dynamics, an interesting question is how policy-makers could exploit the insights of the Social Bubble Hypothesis. It would also be interesting to apply the Social Bubble Hypothesis to science and address the question whether we can identify bubble-dynamics in scientific research and fields. If so, it could be productive to examine how scientific bubbles, which accelerate scientific progress, could be incubated.

The aim of the thesis is to advance our understanding of the nature and structure of technological innovation. Obviously, the development of an exhaustive theory of technological change lies beyond the scope of a PhD thesis. As the question concerning techno-scientific progress is one of the most important question of our age, what is needed is a highly ambitious and multi-disciplinary approach that—by mobilizing methods and insights from economics, complex systems science, philosophy, sociology, and history—attempts to build a systematic and holistic theory of techno-economic progress. Although it does not provide an exhaustive answer—if this is even possible—this thesis should have at least revealed the civilizational importance of the question concerning the nature and structure of technological innovation.