Synchronous failure: the emerging causal architecture of global crisis

Thomas Homer-Dixon1, Brian Walker2, Reinette Biggs3,4, Anne-Sophie Crépin3,5, Carl Folke3,5, Eric F. Lambin6,7, Garry D. Peterson1, Johan Rockström1, Marten Scheffer8, Will Steffen3,9 and Max Troell1,3

ABSTRACT. Recent global crises reveal an emerging pattern of causation that could increasingly characterize the birth and progress of future global crises. A conceptual framework identifies this pattern’s deep causes, intermediate processes, and ultimate outcomes. The framework shows how multiple stresses can interact within a single social-ecological system to cause a shift in that system’s behavior, how simultaneous shifts of this kind in several largely discrete social-ecological systems can interact to cause a far larger intersystemic crisis, and how such a larger crisis can then rapidly propagate across multiple system boundaries to the global scale. Case studies of the 2008-2009 financial-energy and food-energy crises illustrate the framework. Suggestions are offered for future research to explore further the framework’s propositions.

Key Words: climate change; conventional oil; financial system; global crisis; grain supply; social-ecological system

INTRODUCTION

It has recently been proposed that identifiable boundaries mark the safe limits of human alteration of planetary biophysical systems such as nitrogen and carbon cycles (Rockström et al. 2009a, b, Steffen et al. 2015). By this view, exceeding specific levels of key variables in these systems significantly raises the likelihood of a regime shift, which would be a sharp, nonlinear jump or “critical transition” to an alternate state (Scheffer 2009, Barnosky et al. 2012). Should such a transition occur at an Earth-system level, it could affect vital social and economic systems and quickly degrade humanity’s condition. For instance, a critical transition in Earth’s climate could cause a sudden drop in world food output that then produces, in the absence of an adequate response, chronic subnational violence, state failure, and broad international conflict.

However, this is a grossly simplistic account of how a catastrophic social-ecological crisis of global scope might occur. In the real world, any such crisis will have an intricate causal, spatial, and temporal structure. For example, rather than a single critical transition at the planetary scale, smaller crises originating within particular systems or geographical regions might propagate across system boundaries, connect together, and then expand into a global crisis (Lee and Preston 2012).

We argue that recent global crises, especially several that occurred simultaneously in 2008-2009, reveal an emerging pattern or architecture of causation that will increasingly characterize the birth and progress of crises in the future. In a conceptual framework that consists of a set of linked propositions, we identify the deep causes, intermediate processes, and ultimate outcomes of this pattern, which we call “synchronous failure” (Homer-Dixon 2006, Kent 2011).

Scholars and commentators have recently begun to highlight how multiple, simultaneous, and interacting global stresses, such as demographic pressure, climate change, resource scarcities, and financial instability, are increasing global systemic risk (Beddington 2009, OECD 2011, WEF 2012, Helbing 2013, PaIMlin and Armstrong 2015). They often describe the situation that humanity faces now and in coming decades as a “perfect storm” of simultaneous crises (Sample 2009, Ahmed 2011, Ehrlich and Ehrlich 2013, Morgan 2013). Although evocative, this phrase implies that the crises align solely by chance. We argue rather that their simultaneity is a manifestation of an underlying causal pattern that is becoming more prevalent, and we elaborate a conceptual framework that provisionally describes this pattern.

By providing a diagnosis of the emerging landscape of global risk, we hope to deepen scientific inquiry into humanity’s future challenges. The integrated framework proposed here shows how multiple stresses can interact within a single social-ecological system to cause a shift in the system’s behavior, how simultaneous shifts of this kind in several largely discrete social-ecological systems can interact to cause a far larger intersystemic crisis, and how such a larger crisis can then rapidly propagate across multiple system boundaries to the global scale.

We believe we are describing a fundamentally new situation with deep ethical as well as practical implications. Crises have, of course, regularly punctuated human affairs throughout our species’ history. However, when compared with the emerging pattern of crisis we discuss here, past crises, we argue, were generally less global in scope. Those that became truly global, such as the Great Depression and World War II, were rare and thus notable partly because of their global character. Perhaps more importantly, past crises appear to have been generally less intersystemic in their causes and consequences (Biggs et al. 2011).

Because past crises were less globally extensive and intersystemic, substantial resources external to the affected societies remained available for repair of these societies.[1] Partly as a result, affected societies could learn, innovate, and then perhaps fundamentally transform themselves and their institutions in the process of

1 Balsilie School of International Affairs, University of Waterloo, Canada, 2CSIRO Land and Water, Australia, 3Stockholm Resilience Centre, Stockholm University, Sweden, 4Centre for Studies in Complexity, Stellenbosch University, South Africa, 5Beijer Institute of Ecological Economics, Royal Swedish Academy of Sciences, Sweden, 6Earth and Life Institute, University of Louvain, Belgium, 7School of Earth, Energy & Environmental Sciences and Woods Institute for the Environment, Stanford University, United States, 8Environmental Sciences, Aquatic Ecology and Water Quality Management, Wageningen Agricultural University, Netherlands, 9Fenner School of Environment and Society, Australian National University, Australia
rebuilding anew, though sometimes learning was limited and crisis eventually returned. Now, we argue, in an era of global change often called the Anthropocene (Crutzen and Stoermer 2000, Crutzen 2002, Steffen et al. 2007), such second chances are becoming rarer. The global economy has expanded nearly 20-fold since the 1950s, as measured by total global GDP, and inputs of resources from natural systems and outputs of waste back into those systems have increased about 7-fold (Krausmann et al. 2009; J. B. DeLong, unpublished manuscript, http://holtz.org/Library/Social%20Science/Economics/Estimating%20World%20GDP%20by%20DeLong/Estimating%20World%20GDP.htm). Many of these natural systems, including forests, fisheries, hydrological cycles, and atmospheric and terrestrial sinks are consequently under enormous strain, and some, such as Earth’s climate, are exhibiting a higher frequency of extreme behavior (Hansen et al. 2012, IPCC 2012). During the same period, the revolution in information technologies, the quintupling of global trade, and the homogenization of human institutions, culture, and technologies have produced a sharp increase in the connectivity and the speed of operation of human social and economic systems (Chase-Dunn et al. 2000, Young et al. 2006).

The combined result of these changes has been the emergence for the first time in human history of a single, tightly coupled human social-ecological system of planetary scope. Tight coupling means a major crisis is more likely to implicate most if not all of this system, so affected societies will be less likely to have access to largely unaffected societies and regions from which they can draw resources, capital, and knowledge for repair and rebuilding. In a world where external reserves of resources are limited and second chances are thus increasingly rare, humankind must develop the ability to proactively navigate away from this new kind of crisis—globally extensive and intersystemic—that could otherwise irreversibly degrade the biophysical and economic basis for human prosperity. Unfortunately, humanity’s existing global institutions are ill-equipped to provide for such navigation (Walker et al. 2009). We conclude, therefore, with recommendations for further research on both emerging patterns of global crisis and on pathways for rapid institutional evolution that could substantially reduce the risk of synchronous failure.

CONCEPTUAL FRAMEWORK

In this section, we provide definitions of the core concepts and identify the main components of a conceptual framework that describes synchronous failure. A complex systems ontology informs this framework (for a survey, see Mitchell 2009).

Definitions

We define a “system” as a set of causally connected entities that can be considered as a whole and has sufficiently strong homeostatic mechanisms to persist as an identifiable whole over an extended period of time. The entities making up human social-ecological systems include resources, nonhuman organisms, people (including their cognitive states), organizations, institutions, and technologies (Beddoe et al. 2009).

A social-ecological system’s “coping capacity” is its ability to remain within its current stability domain or otherwise shift noncatastrophically to an alternative stability domain in which the system’s essential features and functions are sustained. Our concept of coping capacity echoes that of “resilience” (Folke et al. 2010, Anderies et al. 2013). Our use of coping capacity, however, is intended to emphasize actors’ agency, specifically, the capacity of individuals or groups to choose how, when, and where to respond to systemic stress.

A “stress” is a force that, if unopposed, will move a system away from its current state. A “crisis,” in the context of this paper, is a sudden event or a closely connected series of events within or across social-ecological systems that significantly harms, in a relatively short period of time, the well-being of a large number of people.

This definition of crisis is substrate-neutral, in the sense that a crisis so-defined can occur in an ecological, economic, political, or technological system, or some combination of the foregoing. The definition thus allows researchers to subsume under a single analysis types of events that are conventionally treated separately. The proposed definition stipulates three necessary conditions for crisis that, taken together, are sufficient for its occurrence: suddenness (sharp nonlinearity), impact on a large population, and significant harm to that population within a relatively short period of time. Each of these conditions or properties can be operationalized.

In our stipulation that a crisis must actually cause significant harm, we depart from the conventional usage in which an event that could potentially lead to significant harm is sometimes called a crisis. Such usage is common in discussions of international affairs, as in the label “Cuban Missile Crisis.” By defining crisis as a phenomenon that actually causes harm, we make its identification and measurement less dependent on assessment of people’s subjective psychological states of, for instance, fear, anxiety, and perception of threat.

The definition also assumes that the harm in question is a direct result of the crisis and therefore occurs relatively rapidly. It thus does not preclude the possibility that a crisis can precipitate widely beneficial changes in psychological states, social structures, and general well-being over a longer period of time.

Finally, we do not stipulate that unexpectedness, or perceived low probability, is a necessary condition for crisis, as do some widely cited definitions.[2] First, events generally understood as crises are often broadly anticipated. If Israel attacks Iran to destroy its nuclear facilities in 2016, the event and its immediate consequences will be called a crisis, even though they have been anticipated for years. Second, nearly all crises are anticipated by someone, even if only by chance or luck. Therefore, unexpectedness is not a binary variable, that is, it is not true that sudden harmful events are either expected or unexpected. Because unexpectedness is a continuous variable, stipulating it as a necessary condition would require stipulating what proportion of the population the crisis must truly surprise, a significant and unnecessary complication.

Researchers could use our definition to generate a list of crises across recent history, say, the last century. They could then examine the list to establish trends regarding crisis frequency, type, and severity and to determine whether new patterns of crisis are emerging.
Deep causes of synchronous failure

We argue that compared to past crises, future crises will increasingly arise from the conjunction of three underlying, long-term, and causally linked global trends. The first is the dramatic increase in the scale of human economic activity in relation to Earth’s natural resources and systems. Human-induced changes in natural systems now often rival or exceed changes arising from nonhuman processes (Steffen et al. 2007). The second trend is the rapidly rising density, capacity, and transmission speed of the connections carrying material, energy, and information among the components of human technological, economic, and social systems (Helbing 2013). This increased connectivity reduces the isolation of these systems’ components from each other and thereby increases the functional size of the overall systems of which they are a part. The third trend is the increasing homogeneity, or declining diversity, of human cultures, institutions, practices, and technologies (Boli and Thomas 1997, Meyer 2000, Young et al. 2006), including technologies that exploit ecosystem services, such as agriculture and aquaculture. Although market competition in the global economy can promote diversity, positive network externalities, winner-take-all dynamics, and efforts by firms to achieve efficiencies and economies of scale across enormous markets encourage process homogeneity and a concomitant pruning of redundancy and system slack (Levitt 1983, Arthur 1994, Frank and Cook 1996).

The second and third of these trends are reciprocally related, that is, they are both causes and consequences of each other, although not exclusively so. Greater connectivity facilitates homogenization, while homogenization encourages greater connectivity.

The three global trends contribute both separately and in combination to conditions favoring synchronous failure in three major ways. First, they generate multiple simultaneous stresses affecting human societies. These stresses build their force slowly yet are potentially very powerful over time. For instance, the first of the three above trends, the sharply rising scale of human economic activity in relation to natural resources and systems, is causing greater scarcity of some critical resources such as conventional oil (Sorrell et al. 2012, IEA 2013, Höök et al. 2014), where this scarcity is gauged by the amount of energy needed to extract and process an additional increment of final output (Davidson et al. 2014). It is also contributing to higher atmospheric concentrations of greenhouse gases, which are boosting the incidence of extreme climate events such as heat waves (Hansen et al. 2012, IPCC 2012). Additionally, it is producing severe disruption of many natural systems that are vital to human well-being, including the majority of Earth’s fisheries and large tracts of its grasslands and forests. For some of these systems, such as coral reefs, disruption is approaching or even exceeding the systems’ homeostatic capacity to maintain their integrity and identity (Bellwood et al. 2004, Hughes et al. 2010).

Combinations of the three global trends can also produce stress. For example, climate change and the rising energetic cost of oil, both a result of the first trend, are encouraging a transition from energetically dense carbon-based fuels to alternative energy sources that are, on average, energetically costlier and less dense (Heinberg 2009). However, enabling and sustaining the second trend, the rising connectivity and in turn complexity of human societies, requires ever-larger inputs of high-quality energy (Odum 1988, Tainter et al. 2003). The fundamental contradiction between the two trends could cause enormous economic and social disruption this century (Morgan 2013).

Second, the three global trends contribute to conditions favoring synchronous failure by increasing the risk of large and abrupt systemic disruption and by helping such disruptions propagate farther and faster through global networks. For instance, the increasing density, capacity, and transmission speed of connections among system components creates tight coupling among these components, which raises the risk of surprising and harmful interactions among them (Perrow 1999) and of localized failures spreading quickly to other, distant system components (Buldyrev et al. 2010, Harmon et al. 2010, Bashan et al. 2013, Helbing 2013). Research on ecological and other complex networks suggests that the combination of rising connectivity and homogeneity makes a system less adaptive and, ultimately, more vulnerable to a critical transition, that is, to a wholesale systemic shift or crash (Bodin and Norberg 2005, Scheffer et al. 2012, Lever et al. 2014).

However, the same features can also produce benefits. For instance, greater connectivity can aid repair of local failures by facilitating the flow of inputs from nearby components (Biggs et al. 2012, Scheffer et al. 2012). Also, the combination of greater connectivity and homogeneity in human social systems could set the stage for beneficial critical transitions, such as a planetary shift toward more environmentally sustainable values, institutions, and economic practices.

The above development, an increased propensity to large disruptions that propagate farther and faster through global networks, may be evidence of a more generalized trend of direct relevance to our argument here: greater synchronization in global systems. Complex systems often exhibit synchronization (Strogatz 2003). This is an “emergent” phenomenon, that is, a novel property of a whole system that arises from the interactions of its component parts. Although the underlying causes of synchronization are still not fully understood, it appears likely to arise when a dense network of links among largely homogenous system components carries signals that create positive feedbacks among these components. The three global trends identified above cause today’s global systems to increasingly manifest exactly the features of dense connectivity, long-distance telecommunications, and component homogeneity that seem to encourage synchronization (Biggs et al. 2011). We will highlight the apparent role of the global energy system in synchronizing societal crisis.

Finally, the three global trends can contribute to conditions favoring synchronous failure by reducing the coping capacity of societies. Greater connectivity within and among technological systems and their associated social systems often makes these systems’ internal structures and mechanisms more opaque to system managers, boosts the cognitive and decision-making load borne by those managers (Barrett et al. 2004, HIMSS 2009), and generally makes system prediction and control harder. Opaqueness, managerial overload, and low control are increasingly apparent in systems as diverse as electrical grids and the global financial and supply-chain networks (Korowicz 2012, Lee and Preston 2012, Toltefson 2013).

Processes of synchronous failure

The deep causes we identify above manifest themselves in three crisis processes that, together, are core characteristics of
synchronous failure. Each of these processes usually operates in conjunction with at least one of the other two; that is, each is rarely seen in isolation. However, each is archetypal and as such, we believe, deserves a distinct label.

**Long fuse big bang**
The first process, which we call “long fuse big bang” (or LFBB), arises from the slow accumulation of a stress or stresses in a social-ecological system (the burning of the long fuse) that ultimately produces a rapid and sharply nonlinear shift in the system’s behavior (the big bang), perhaps to a new system configuration. The shift occurs when the level of stress exceeds the system’s coping capacity. In other words, it occurs when the ratio of stress to the system’s coping capacity exceeds one, assuming that valid, reliable, and independent instruments to measure both stress and coping capacity are available. We call this situation “overload.”

Our conceptualization of the LFBB archetype derives from two distinct yet complementary sources: political science theory on societal crisis and complex systems theory on nonlinear change. Political scientists Karl Deutsch (1954), Samuel Huntington (1968), and Alexander Motyl (2001), among others, have proposed overload-breakdown theories of social and political crisis. Huntington, for instance, argues that societies are vulnerable to instability when their level of political participation exceeds their level of political institutionalization. Such theories, in turn, trace their origin to at least three distinct lines of previous thought (Homer-Dixon 2006): the functionalist sociology of Émile Durkheim (1951) and Talcott Parsons (1951), the systems theory and cybernetics pioneered by Ludwig von Bertalanffy (1950) and Norbert Wiener (1961), and the information-processing and computational theories of cognitive scientists and organizational theorists such as Herbert Simon (1983).

The other source of our conceptualization of LFBB, complex systems theory, proposes that sharp nonlinear shifts in system behaviour can arise in two distinct ways. In the first, change in the values of one or more state variables causes the system to jump suddenly from one place to another in a multidimensional state space; this system “flip” or critical transition can be modeled as a bifurcation on a catastrophe manifold (Scheffer 2009). In the second type, potential energy that has slowly accumulated within the system is released in an “avalanche” event. In systems that exhibit self-organized criticality, a power law describes the frequency distribution of avalanche size (Bak 1996). Very large events, that is, avalanches in the tail of the distribution, are possible.

**Simultaneous stresses**
The second process of synchronous failure, which we call the “simultaneous stresses” (SS), arises when two or more stresses combine within a single social-ecological system. Correctly describing the relationship among these multiple stresses is important, and here two key questions arise. First, are the stresses causally linked or causally independent? Second, do the stresses, in their combined effect, have a multiplicative or additive relationship?

In regard to the first question, stresses developing within the tightly coupled social-ecological systems that increasingly characterize today’s world will rarely if ever be fully causally independent of each other. However, the causal relations between them can take a variety of forms, from linear unidirectional causation at one extreme to highly nonlinear reciprocal (or feedback) causation at the other extreme. Also, causal links between stresses can occur along the full range of temporal stages in the stresses’ development. Any adequate account of multiple stresses must provide detail on such complexities.

In regard to the second question, a multiplicative relationship among stresses implies that their combined effect, i.e., total stress, is not a straight-forward sum of their individual contributions to that effect; the relationship among the stresses is therefore synergistic. Also, a multiplicative relationship among stresses implies a logical “and” relationship among them, which means each stress is posited as necessary for the result to occur. In other words, overload can happen only when stresses X and Y occur together in sufficient combined strength to exceed the system’s coping capacity. The outbreak of a vector-borne disease in a population, for example, requires both a pathogen and a vector. Each is necessary for the outbreak to occur, so neither can cause the outbreak by itself. An additive relationship, on the other hand, implies a logical “or” relationship, which means any one of the identified stresses by itself could be sufficient for the outcome. In this case, overload happens when X and/or Y occurs in sufficient strength to exceed the system’s coping capacity (Mahoney 2008, Mahoney et al. 2009).

The actual relationship between stresses in a given situation must be determined through empirical investigation. It is quite possible, indeed, that the best representation will involve a combination of the two types of relationships. For instance, overload may be found to happen when (X or Y) and Z occur together. Also, researchers may determine that the overload outcome is “equifinal,” which means it could arise along multiple discrete causal pathways. In this case, researchers should ideally stipulate each pathway’s sufficient set of causes.

The product of the exercise will resemble the kind of fault-tree model common in risk analysis (Lee et al. 1985).

**Ramifying cascade**
The third process archetype, which we call the “ramifying cascade” (RC), arises when a sudden and severe perturbation of one node in a tightly coupled network propagates through the network’s links to adjacent nodes, thereby producing knock-on effects at some distance from the original perturbation. We argued above that steadily greater connectivity within human technological, economic, and social networks and between these human networks and natural systems has increased the frequency and severity of unexpected system interactions and of long-range cascading system failures (Adger et al. 2009, Galaz et al. 2011, Korowicz 2012).

Each of our process archetypes highlights a distinct and critical property of the emerging pattern of social-ecological crises we call synchronous failure. The LFBB archetype emphasizes how a stress on a system can cause a sudden nonlinearity, in which the rate of system change shifts abruptly from slow to fast. The SS archetype highlights how multiple stresses operating simultaneously can combine in their total impact, often synergistically. Finally, the RC archetype highlights how disruptions or shocks arising from sudden nonlinearities can propagate rapidly through modern tightly coupled social-ecological networks.
Outcomes of synchronous failure

We argue that the above causes and processes are increasingly producing crises with a distinct character compared to past crises. Specifically, synchronous failure, as we call this emerging type of crisis, is more biophysical in origin, more intersystemic in manifestation, more global in scope, and more rapid in development. Together, these four properties increase the risk that future crises will involve irreversible system flips on human timescales that have enormous repercussions for humankind.

In Figure 1, we bring together the various elements of our conceptual framework. The figure distinguishes between two temporal stages. In stage 1, which integrates the SS and LFBB process archetypes, slow processes operate within largely discrete systems. Multiple slowly developing stresses combine to cause overload in these systems. With the "X" symbols between the converging arrows, we specify that the relationships between the stresses in each system are multiplicative; as discussed above, however, these relationships could also be additive.

The operation of proximate triggers marks the transition between stage 1 and stage 2; these triggers cause the overload in each system to generate discrete systemic crises. The simultaneity or near-simultaneity of these crises, that is, the synchronization of their crisis behavior, arises from the systems' causal interaction across their intersystemic boundary.

In stage 2, which integrates the SS and RC archetypes, fast processes operate across multiple systems. Here the simultaneous but previously largely discrete systemic crises combine, again synergistically, to produce a single multisystemic crisis that cascades outward through a tightly coupled global network.

The SS process archetype appears in both stages. In the stage 1, it is slow-acting stresses that are multiple; whereas in stage 2, it is fast-acting crises that are multiple. The RC archetype by definition involves fast processes, whereas LFBB involves both slow and fast processes—at first slow and then fast. Our framework implies that the transition between slow and fast system behavior that produces a crisis requires something like a LFBB process. The specific causal mechanisms that produce this nonlinear response will differ from system to system.

Some might object that the conceptual framework offered here is both too deterministic and too pessimistic. In the past, these critics might argue, multiple stresses and even crises have often spurred innovation, reorganization, and adaptation rather than catastrophic system failure. Now humanity's situation offers even more reason for optimism: scientific and technological progress has combined with the globalization of ideas and communication systems to greatly boost humankind's ability to anticipate, understand, avoid, and respond constructively to crises. Our framework thus underestimates human societies' longer term coping or adaptive capacity. Short-term overload and crisis will lead in time to creativity and adaptation that leave societies better off than they were before (Wilkinson 1973, Simon 1998).

We acknowledge this objection but argue in response that humanity's predicament this century is largely novel and is best characterized as a race between the rapidly increasing severity and complexity of its problems and its improving but nonetheless
uncertain ability to anticipate, proactively solve, and constructively respond to these problems (Homer-Dixon 2000). In 2015, we cannot be sure of this race’s outcome. Undoubtedly, crisis can create opportunities for innovation. However, human history offers abundant evidence that these opportunities are often not exploited or that, if they are, the resulting innovation is either insufficient or of the wrong kind to produce a long-term sustainable outcome (Homer-Dixon 2000). Also, many of the most severe problems that humankind now faces, such as climate change, biodiversity loss, and nutrient pollution involve the provision of global public goods (Kaul et al. 2003, Barrett 2007). Because property rights for these goods are nonexistent, unclear, or shared, the market institutions that have become the primary mechanisms of human innovation often cannot generate an adequate response (Cornes and Sandler 1996).

Humanity will respond more effectively to its evolving challenges if it better understands these challenges. We propose our conceptual framework with this goal in mind. It consists of a series of interlinked propositions about the emerging properties of global crisis that, we believe, deserve extended close examination.

ILLUSTRATIONS

Illustration 1: the 2008 financial-energy crisis

All of our framework’s components are visible in the crisis that erupted in financial markets and the global economy in 2008. Most people recall that ongoing troubles in the U.S. subprime housing market metastasized into a global financial crisis, precipitated, especially, by the late-September collapse of the investment bank Lehman Brothers. However, another event of global scope occurred that year too: a rapid run-up and then collapse in oil prices. The price of light crude oil rose from US$90 to nearly $150 dollars a barrel between January and June and then plummeted to less than $50 by the end of the year. Although it may seem that the financial crisis and the energy-price shock were causally independent phenomena, we argue the two events were connected in their deep causes, proximate triggers, and intertwined consequences, as illustrated in Figure 2. We therefore label the event a “financial-energy” crisis. It exhibited all four properties that we believe will increasingly distinguish future crises: it was global in scope; it developed extraordinarily rapidly; it was clearly intersystemic, because it emerged from the interaction of humanity’s financial and energy systems; and it was profoundly biophysical, because it was rooted, in significant part, in humanity’s use of natural resources.

The global energy system is represented at the bottom of the figure. Between the mid-1990s and 2008, several slowly building stresses combined synergistically within that part of the system involving liquid transportation fuels. Global demand for liquid fuels increased one to two percent a year, driven especially by the giant economies of China and India as they continued energy-intensive development and expanded their transportation fleets. At the same time, a rising proportion of the planet’s conventional oil production came from fields that had passed peak output and were exhibiting year-over-year production declines (Sorrell et al. 2012, IEA 2013, Höök et al. 2014). Finally, the energetic cost of
finding and producing new sources of oil continued to rise steadily, which meant that the energy return on investment (EROI) fell for both conventional and nonconventional sources (Guilford et al. 2011, Fournier et al. 2013, Murphy 2014, Hall et al. 2014).[5] These stresses combined to sharply reduce slack in the global oil market, making it more tightly coupled and reducing the elasticity of supply for liquids (Murray and King 2012).

The global financial system is represented on the top of the figure. In United States in the 1980s, economic policy makers broke with the prevailing Keynesian orthodoxy and adopted monetarism, largely in response to the country's experience with stagflation during the previous decade. Federal Reserve Chairman Paul Volcker used sharply contractionist interest-rate policies to bring inflation under control. Once inflation subsided, economic policy makers around the world responded to chronically insufficient demand, and the labor and capital unemployment it produced, by encouraging growth in the supply of credit in the economies under their direction. An antiregulation economic ideology, which complemented the monetarist shift, took hold in key policy circles, especially in the United States.

The result was a two-decades-long torrent of liquidity into the global economy that produced rapid economic growth, especially in Asia (Reinhart 2012). However, the flood of liquidity also helped produce a succession of speculative bubbles, in East Asia in the 1990s, in the American information technology sector in the late 1990s and early 2000s, in the housing sectors in the U.S. and some parts of Europe in the late 2000s (Justiniano et al. 2015), and in natural resource markets in the late 2000s.

In this context, a number of economic stresses combined synergistically during the decade prior to 2008. Easy credit and deregulation encouraged the practice of securitizing debt and risk, which spread across all sectors of advanced economies and widened the gap between creditors and the underlying assets securing their credit. This widening gap, in turn, eroded the ability of, and the incentive for, creditors and investors to judge asset risk accurately. Meanwhile, equity traders used increasingly sophisticated computational and mathematical technologies to generate securitized assets and estimate their risk, along with faster communication technologies to trade these assets globally. The rising complexity of these assets and the increased density and speed of global interconnections between investors, investments, and markets made the behavior of the overall economic system progressively more opaque. Finally, changes in the architecture of the global financial network tightened its coupling and reduced modularity, increased the concentration of assets and network connections in a small number of immense financial institutions, and depressed the diversity of liabilities and assets on banks' balance sheets (Haldane and May 2011, Forbes 2012).

As a result of these trends in combination, policymakers, regulators, economic commentators, and investors increasingly underestimated systemic risk, that is, risk at the level of the global economy as a whole. The prevailing antiregulation ideology gave key rating agencies such as Standard & Poor's and Moody's latitude to assign triple-A ratings to what were in reality extremely low-quality assets.

The left side of Figure 2 represents the SS and LFBB properties of both the energy and financial systems. These systems were distinct, but they were not fully isolated from each other. Indeed, even before the stresses identified above were severe inside both systems, links of two kinds had developed across their shared boundary.

First, there were causal connections from the economic system to the energy system. The liquidity-driven economic growth from the late 1980s to the mid-2000s led directly to surging energy and especially oil consumption. By 2008, speculators with access to cheap credit and advanced trading technologies were multiplying the impact of the underlying absence of slack in the global oil market (Juvenal and Petrella 2012).

Second, there were connections in the reverse direction—from the energy system to the economic system. From the 1986 to the early 2000s, the energy available in oil was inexpensive by historical standards. It was cheaper, in real terms, than during the entire period from 1973 to 1986. Inexpensive energy along with technical advances such as containerization meant inexpensive transportation, and inexpensive transportation encouraged the development of globe-spanning networks of producers and suppliers. The result was a rapid increase in the material connectivity of the global economy that paralleled the soaring information connectivity arising from new information technologies (Chase-Dunn et al. 2000).

These bidirectional interactions, we propose, helped synchronize the behavior of global economic and energy systems, so that by mid-2008 both simultaneously reached critical junctures and exhibited extreme susceptibility to crisis. Within each system, accumulating stresses had overloaded normal coping mechanisms, in particular the usual regulating and stabilizing effects of markets.

For instance, in the financial system, U.S. banks could not write off fast enough the rapid accumulation of nonperforming securitized debt tied to subprime mortgages. Simultaneously, in the global energy system in the first half of 2008, rising oil prices encouraged heavy speculation in oil futures, driving prices even higher. This was a positive or self-reinforcing feedback. Normally, price signals in oil markets stimulate negative-feedback behavior: higher prices bring to market greater oil supply just as they reduce demand by encouraging conservation; both outcomes, in turn, lower prices.

Researchers have not reached a consensus as to the proximate trigger of the ensuing global crisis. However, the second stage of this episode of synchronous failure may have begun in the vast suburban and exurban tracts of housing in California, Nevada, Arizona, and Florida that were the prime sites of the mid-2000s U.S. real-estate bubble. Development of these tracts made commercial sense only when gasoline was cheap: many residents had precarious finances at the best of times and nearly all had long commutes to work, often two hours or more. As the U.S. price of gasoline passed the psychologically critical threshold of $4 a gallon, as it did in early June of 2008, a sharply higher fraction of residents and of potential buyers found houses in these zones unaffordable. Defaults on mortgages on the houses soared, which meant that the value of securities backed by the mortgages plummeted.[6]

The months of June through September of 2008 thus saw the energy and financial crises, until then largely discrete, combine to spark a global, multisystemic crisis. Markets did not expect
Lehman Brothers’ late-September collapse, and they interpreted the event as indicating the U.S. Federal Reserve could not manage the rising tide of defaults (Swedberg 2010). Securitized assets on the balance sheets of banks and corporations around the world suddenly had unknown value. The interdependence of bank, hedge fund, and corporate balance sheets tightened the coupling of global financial markets, so uncertainty propagated rapidly through global markets. Banks and firms, not confident of the financial viability of counterparties whose balance sheets they could no longer evaluate, simply stopped doing business with each other. Global trade collapsed (Levchenko et al. 2009), and every major economic region in the world began to contract simultaneously. Falling demand drove oil prices down nearly one hundred dollars a barrel in three months.

Only history’s largest coordinated central-bank intervention, involving the injection of trillions of dollars of additional liquidity and outright nationalization of hundreds of failing banks and industries, halted the crisis in late 2008 and early 2009. However, many of the underlying stresses that led to the crisis, including the rising energetic costs of transportation fuel and the generally weakly regulation of financial markets, continue to operate today.

Illustration 2: the 2008 food-energy crisis

Between 2006 and mid-2008, the average global price of food nearly doubled, and in the first half of 2008 the price of grain in particular shot upwards. Because poor people usually spend a large portion of household income on food, these sharply higher prices hurt tens of millions of people around the world, causing a temporary surge in malnutrition in poor countries and widespread decline in well-being, particularly in urban zones and among children (Ivanic and Martin 2008, Tiwari and Zaman 2010).

Researchers generally explain this event as the combined result of various weather-related shocks to agricultural trade, the promotion of biofuels by rich countries, and the surge in the cost of energy inputs to agriculture due to higher oil prices (Headey and Fan 2008, Headey et al. 2010, Lagi et al. 2011, Wright 2011). Export bans by a number of key food-producing countries then amplified the spike.

This research rightly emphasizes that many factors interacted to cause the price surge and that the global energy system played a central role. However, for the most part it does not distinguish between long-term stresses and more temporally proximate causes. Nor, generally, does it explain the mechanisms underlying the food system’s sharply nonlinear response. Finally, it insufficiently highlights the crisis’s deeply intersystemic character.

We capture these properties in Figure 3, which represents the 2008 food crisis in terms of our framework. The crisis and more recent high-amplitude swings in food prices suggest that humanity’s food system is under extreme pressure and may experience much larger crises in the future (Berry et al. 2012).
Over thousands of years, humankind extensively transformed Earth’s biosphere as agriculture gradually expanded and diversified. In the past century, agriculture’s pace of change became much more rapid and its biospheric impacts even more pronounced. Traditionally, diverse agricultural systems of shifting cultivation that recycled nutrients through fallow practices or agroforestry gave way to ecologically far simpler systems relying heavily on fossil fuels for mechanized labor, irrigation, and inputs of fertilizer and pesticide. These new systems originated in Europe and then spread to the Americas, Asia, and finally parts of Africa. They produce cheap food that has enormously improved human well-being. Nonetheless, because of poverty, about 1 billion people remain malnourished today.

In recent decades, four simultaneous and sometimes synergistically interacting stresses have been building within this global food system. The first is steadily diminishing availability of new, high-quality agricultural land. Agriculture uses about 35% of the world’s land. Further expansion is possible, but only a few countries, such as Angola, Brazil, Congo, Mozambique, Russia, Zambia, and Tanzania, have enough unused land of sufficient quality to sustain any substantial increase. There is substantially less potential additional cropland than is generally assumed once ecological and socioeconomic constraints and trade-offs are taken into account (Lambin et al. 2013).

The second stress is steadily declining marginal returns to agricultural intensification in many parts of the world, that is, to incremental additions of inputs such as fertilizer, machinery, and irrigation water (Fuglie 2010). In rich countries, intensification’s marginal benefits are already low: it can boost agricultural yields, but sustaining past rates of output growth through intensification alone usually produces external costs, such as greater water pollution and loss of ecosystem services, and these costs can significantly offset any benefits (Matson and Vitousek 2006, Bennett et al. 2014, Rist et al. 2014). In poorer regions where big gains from intensification are more achievable, social, infrastructure, and institutional problems often hinder such a strategy. In the absence of greater investment in inputs themselves—use of more fertilizer and tractors of the same type, for instance—better technologies can raise input productivity. A kilogram of fertilizer will produce much more food if technology can tell farmers which parts of their land need it most. Indeed, in the last decade, higher input productivity accounted for three-quarters of growth in the world’s food output (Fuglie 2010, Fuglie et al. 2012), but productivity improvements appear to be slackening for some cereals, especially wheat and rice (Ray et al. 2012, Grassini et al. 2013).

Third, climate change has begun to affect food output by increasing, in particular, the frequency and severity of extreme weather events such as drought (Lobell and Field 2007); this impact is expected to become much more pronounced in coming decades, requiring ever-larger investments in new agricultural technologies and infrastructure (Battisti and Naylor 2009, Ackerman and Stanton 2013, Dai 2013).

These first three stresses operate, essentially, on the supply side of the global food system. Although they have not halted output growth (Fuglie et al. 2012), they have constrained this growth and in some cases offset its economic and social benefits. Meanwhile, world food demand is rising relentlessly, constituting a fourth stress. Demand is rising partly because the human population is still growing by about 70 million people a year and partly because this population’s steadily higher average income has increased meat consumption. Greater meat consumption, in turn, drives overall food consumption up the trophic hierarchy, requiring greater total food production (Bonhommeau et al. 2013). Higher demand for food generates opportunity for profit from investment in agriculture, but the three supply-side stresses described above can decrease the attractiveness of such investment. Low investment further constrains future growth in global food output and raises the potential for even larger food-price spikes in coming decades.

The upper left half of Figure 3 represents the global food system, showing the four above-described simultaneous stresses. Together, the stresses exhibited SS and LFBB dynamics. In the years prior to 2007-2008, several of them, specifically, diminished land availability, declining marginal returns to intensification, and rising food demand, started to overload the global food system’s coping capacity, including its market mechanisms and the international institutions tasked with responding to food shortfalls, such as the World Food Program (IDC 2008, WFP 2009). These institutions are supposed to act as stabilizing (negative) feedbacks in the global food system, but their effectiveness is doubtful because of their limited capacity and operational scope (Walker et al. 2009). The bottom half of the diagram again portrays the global energy system with the three main stresses as described previously: increasing global demand for conventional oil, the decline of mature oil fields, and rising energetic cost of finding and producing a marginal barrel of oil.

Over the years prior to the 2008 crisis, connections developed in both directions across the shared boundary between the food and energy systems, once again contributing, we propose, to their synchronized crisis behavior. Cheap oil in the 1980s and 1990s encouraged farmers to increase energy inputs to food production, through greater mechanization and irrigation and greater use of fertilizer. Rising oil prices in the 2000s then directly boosted the cost of food (Headey and Fan 2008, Baffes and Dennis 2013). Rising prices also encouraged farmers to use their land to grow biofuels that substitute for petroleum-derived transportation fuel. Because cropland can now grow either food or feedstock for biofuel, human agricultural and energy systems are bidirectionally coupled (Hertel and Beckman 2011, Searchinger and Heimlich 2015). Products of the energy system are an input to the agricultural system, and products of the agricultural system are an input into the energy system. Although analysts were aware of the above factors prior to the 2008 food crisis, the event nevertheless surprised most experts. Few anticipated how, in the context of the four stresses on the food system described here, low carryover stocks of grain from 2007 amplified the sensitivity of the global food system to various proximate shocks, such as an extended Australian drought (Wright 2011).

The world’s immediate responses made the problem worse, partly because the dynamics of the crisis were so poorly grasped. In particular, some food-producing countries, including Brazil, India, and Vietnam, banned exports, further driving up prices (Mitra and Josling 2009). Also, many countries exhibited little short-run resilience in the face of food shortages and price
increases, having previously reduced their national stores of grain on the assumption that efficient international food markets made large-scale food crises improbable.

The global financial system played a role too, although it is unclear how much speculation in futures markets affected the crisis. Some research shows evidence of a strong impact, whereas other research has found little (Gilbert 2010, Lagi et al. 2011, Auierich et al. 2013). Nevertheless, it is certainly true that financial deregulation and the development of new classes of derivatives encouraged index-based investment in food markets that were previously used almost exclusively to hedge producers’ risk.

The 2008 food crisis had vast multisystem impacts. We have mentioned its implications for the well-being of families, individuals, and especially children around the world. The surge in food prices also destabilized many political systems: food riots and violence broke out in dozens of poor countries, including Bangladesh, Burkina Faso, Cameroon, Egypt, Indonesia, and Yemen. However, the global food system showed resilience over the medium term. Farmers responded strongly to higher prices: food production in major wheat and maize exporting countries rose by 25% to 30% the following year, and China and India increased their public agricultural spending sharply (Heady et al. 2010). All the same, prices generally remained higher than they had been in the 1990s or 2000s, and they appeared to become more volatile.

These conditions set the stage for further food-system instability. In August 2010, Russian authorities banned grain exports after Russian crops were damaged by severe drought and record heat, a weather phenomenon possibly attributable to climate change. Within weeks, key grain prices surged upwards, and in the fourth quarter of 2010 world food prices overall shot past 2008 levels. By spring 2011, food prices worldwide were averaging 20% above 2008 levels. Higher food prices were likely a proximate trigger of the uprisings that began in Tunisia in December 2010 and convulsed nearly 10 countries in the Middle East and North Africa through the winter and spring of 2011 (Johnstone and Mazo 2011, Werrell and Femla 2013).

CONCLUSION: FUTURE RESEARCH

In this paper, we have elaborated a conceptual framework that shows the emerging causal architecture of global crisis, which we call synchronous failure. This framework identifies the phenomenon’s deep causes, characteristic processes, and common outcomes. We illustrated our framework with accounts of two recent global crises, the 2008 financial-energy crisis and the 2008 food-energy crisis.

In light of our analysis, it is striking that both of these crises occurred in 2008 and that, in both, the global energy system, in particular the global conventional-oil system, played a central role. Indeed, because of the energy system’s role, both crises could be considered together as a single three-system instance of synchronous failure.

As researchers consider how the form and incidence of major crisis is changing, the causal role of the global energy system deserves special attention. Only enormous inputs of inexpensive high-quality energy can create and sustain the unprecedented connectivity and complexity of human civilization, including the connectivity described here among this civilization’s diverse component systems. As a provisional hypothesis, therefore, it seems reasonable to propose that the global energy system helps to synchronize these systems’ behavior and to stimulate simultaneous crises within and across them. Other factors such as global trade and transport systems, the Internet, and simultaneous scarcity of multiple resources (Seppelt et al. 2014) may also play synchronizing roles, but these factors themselves depend on, and are therefore significantly derivative of, massive flows of energy.

Global adjustment to worsening energy scarcity is unlikely to be smooth. The rising energetic cost of energy stimulates boom-bust investment cycles and alternating episodes of glut and scarcity (Jackson and Smith 2014). In late 2014 and early 2015, global oil supply exceeded demand by a relatively small amount, about 1% to 2% on a daily basis, yet international oil prices dropped by 50%. The drop has generated instabilities in the global economy, with key oil-producing economies experiencing a precipitous decline in revenues. Despite the fall in prices, global oil markets still exhibit limited slack and therefore remain vulnerable to geopolitical disruption, arising from, for instance, instability in the Middle East, that could be the proximate cause of larger intersystemic crisis.

More generally, researchers should closely investigate the complex, dense, and evolving causal links among humanity’s energy, food, water, climate, and financial systems, along with the implications of these links for long-term human well-being. Although research programs have explored the relationships between specific pairs of these systems, for instance, the water-food, energy-finance, energy-climate, and climate-food relationships, no research program has explicated the joint and coevolving behavior of all five systems together, especially not in light of the biosphere’s capacity to sustain these systems (Steffen et al. 2007, Galaz et al. 2012). Yet humanity’s fate likely resides at the nexus of these systems.\[3]\n
Our proposed framework could guide a systematic analysis of recent global crises. Such an analysis would help to refine the framework. It could also identify early-warning indicators and appropriate interventions that enhance societal resilience and adaptive governance and, ultimately, reduce the danger of synchronous failure.

\[1\] The collapse of the Soviet Union and its associated empire in the late 1980s and early 1990s is an example of such a geographically and systemically bounded crisis.

\[2\] For instance, Seeger, Sellnow, and Ulmer define an organizational crisis as “a specific, unexpected, and nonroutine event or series of events that create high levels of uncertainty and threaten or are perceived to threaten an organization’s high-priority goals” (Seeger et al. 1998:233).

\[3\] In an additive relationship, one stress can mask the effect of another. For example, high regional concentrations of sulfate aerosols can mask the generalized warming effect of rising atmospheric concentrations of carbon dioxide.

\[4\] Technically, this would be a representation of INUS causation, in which each cause is an insufficient but necessary member of a set of causes that is itself unnecessary but sufficient for the equipinal outcome of system overload (Mahoney 2008, Mahoney et al. 2009).
The EROI trends mentioned here are well-documented (Hall et al. 2014). New technologies like hydrofracking and steam-assisted gravity drainage (SAGD) have brought large quantities of new oil online, especially from unconventional tight-oil and bitumen formations, but this supply is energetically much more expensive than conventional oil (Murphy 2014). Meanwhile, around the world mature conventional oil fields are rapidly declining at about 6% to 7% a year (IEA 2013). Fracking and SAGD technologies may partially compensate for this decline, but at greatly increased energetic cost (Hughes 2013).

Hamilton (2009) writes: “[There was] an interaction effect between the oil shock and the problems in housing. [In] the Los Angeles, Tampa, Pittsburgh, Chicago, and Portland-Vancouver Metropolitan Statistical Areas, house prices in 2007 were likely to rise slightly in the zip codes closest to the central urban areas but fall significantly in zip codes with longer average commuting distances. Foreclosure rates also rose with distance from the center. And certainly to the extent that the oil shock made a direct contribution to lower income and higher unemployment, that would also depress housing demand.”

In the United States, this substitution was reinforced by the 2005 Renewable Fuel Standard, which mandates that producers of transportation fuel blend into their product a specified proportion of biofuels.


Responses to this article can be read online at: http://www.ecologyandsociety.org/issues/responses.php/7681

Acknowledgments:

We wish to acknowledge the substantive contributions of Nick Bostrom, Gary Bowden, Lisa Deutsch, Jerker Lokrantz, Eric Nævdal, Toby Ord, and Anders Sandberg. Reimette Biggs is supported by a Swiss ETH Society in Science Fellowship, and a Mistra grant to the Stockholm Resilience Centre. This article was initiated at the Inconvenient Feedbacks in Global Dynamics workshop, organized in April 2010 in Stockholm as part of the joint research program on “Global dynamics and resilience,” of the Beijer Institute of Ecological Economics of the Royal Swedish Academy of Sciences and the Stockholm Resilience Centre. Economic support from the Anna-Greta och Holger Crafoords foundation and the foundation Till Bröderna Jacob och Marcus Wallenberg’s minne is gratefully acknowledged. Many thanks to Joan Hewer for assisting with research and formatting and to Jerker Lokrantz for preparing the figures.

LITERATURE CITED


Hughes, D. 2013. *Drill, baby, drill: Can unconventional fuels usher in a new era of energy abundance?* Post Carbon Institute, Santa Rosa, California, USA.


