Addressing Economic Crises: The Reference-Class Problem

Xavier De Scheemaekere, Kim Oosterlinck and Ariane Szafarz

Addressing crises raises a sharp reference-class problem in economics. Namely, economic theory lacks an inclusive and consistent classification of crises. This problem stems from the fact that economists tend to adapt their views on crises to recent episodes, and omit averted and potential crises. They thus fail to imagine innovations in the forms crises can take. We identify three main factors that hamper the economic analysis of crises: the lack of imagination, sample biases, and the Peso problem. Using several historical examples, this paper argues that only an ex ante classification of crises derived from a conceptual approach would address the reference-class problem properly. Multiple-solution models offer a promising avenue to solve this problem.

JEL Classifications: B40, G01, G17, N00

Keywords: Economic Crisis, Economic History, Single-Case Probability, Epistemology

CEB Working Paper N° 12/024
September 2012
Addressing Economic Crises: The Reference-Class Problem

Xavier De Scheemaekere
F.R.S.-F.N.R.S. Research Fellow
Université Libre de Bruxelles (ULB), SBS-EM
1050 Brussels, Belgium

Kim Oosterlinck
Université Libre de Bruxelles (ULB), SBS-EM, Centre Emile Bernheim (CEB)
1050 Brussels, Belgium

Ariane Szafarz*
Université Libre de Bruxelles (ULB), SBS-EM, CEB and CERMi
1050 Brussels, Belgium

September 2012

Abstract

Addressing crises raises a sharp reference-class problem in economics. Namely, economic theory lacks an inclusive and consistent classification of crises. This problem stems from the fact that economists tend to adapt their views on crises to recent episodes, and omit averted and potential crises. They thus fail to imagine innovations in the forms crises can take. We identify three main factors that hamper the economic analysis of crises: the lack of imagination, sample biases, and the Peso problem. Using several historical examples, this paper argues that only an ex ante classification of crises derived from a conceptual approach would address the reference-class problem properly. Multiple-solution models offer a promising avenue to solve this problem.

JEL: B40, G01, G17, N00
Keywords: Economic Crisis, Economic History, Single-Case Probability, Epistemology.

* Corresponding author. Address: ULB, SBS-EM, CEB, Av. F.D. Roosevelt, 50, CP114/03 1050 Brussels, Belgium; Phone: +32.2.650.48.65; Fax: +32.2.650.41.88; E-mail: aszafarz@ulb.ac.be
“Imagination is more important than knowledge. For knowledge is limited, whereas imagination embraces the entire world, stimulating progress, giving birth to evolution”.

Albert Einstein

1. Introduction

What are economic crises? How should we classify them? Economists are mostly unable to answer these questions. Instead, they tend to subjectively adapt their views on crises to recent episodes. This paper investigates the epistemological fundaments for the absence of consensus on the classification of economic crises. We conjecture that crises are difficult to define, and hence classify, not only because they are multifaceted, but also because economic theory has so far proven incapable of overcoming the reference-class problem put forward by Reichenbach (1949).

The reference-class problem relates to the ex ante classification of uncertain events. In a probabilistic framework, observations are outcomes driven by an underlying generating distribution. Hence, ex ante classification is necessary for assessing theories using observations. In addition, valid classes must constitute a partition of all imaginable outcomes. Applied to economic crises, the reference-class problem therefore requires imagining and classifying all possible scenarios for future crises. The standard empirical approach to crises is based on statistical thresholds rationalized by the supposed “anomalies” associated with crises. From a probabilistic standpoint, this hardly makes sense because extreme events are naturally consistent with low-probability outcomes. This is, for instance, the motivation for using, say, 5% thresholds in hypothesis testing. If economic crises were nothing else than below- or above-thresholds outcomes driven by otherwise-standard theory, their representation would require no special attention. In contrast, what economists have in mind when speaking about crises are likely specific features that make the underlying model behave differently than during “normal” times. Therefore, threshold-based econometrics is inappropriate to deal with crises. If crises are special, they would rather lead to discarding

---

1 As quoted in “What Life Means to Einstein: An Interview by George Sylvester Viereck” in the Saturday Evening Post (26 October 1929)
asymptotic theory based on standard models. As a result, econometrics alone is unable to properly address and classify crises. There is a strong need for theory, which does not rely on past observations but rather on an \textit{ex ante} conceptual approach.

But do economic crises deserve a specific conceptual approach? According to conventional wisdom, economic crises are dramatic but rare events, each one coming with its own shape. If this is indeed the case, theorizing crises is pointless as science is about general statements. Alternatively, common characteristics make crises a legitimate object for economic study. Addressing this issue requires a sharp \textit{ex ante} definition of economic crises that goes far beyond conventional wisdom. In turn, such a definition is conceivable only once the reference class problem is solved.

The existing literature offers some clues on this issue. While a consensual definition is still missing, many articles provide stylized facts and classify crises. Expectedly, classifications vary across papers because observed crises differ across a wide range of features including the role of fundamentals, the relative importance of banks and securitized debts, the relative importance of private and sovereign debts, the exchange rate regimes and history, and the underlying structure and dynamics (Portes, 1998). As a consequence, studies are hardly comparable. We identify the reference class problem as the main reason for this outcome. Empirical studies classify crises based on past occurrences and therefore ignore both averted and potential crises. A consensual classification tackling the reference class issue should be general enough to encompass all types of imaginable crisis situations.

\footnote{Moreover, crises are supposedly rare and would then create small sample distortions. On top of being highly sensitive to distributional assumptions, empirical results driven from small samples may exhibit large variations with respect to the inclusion/exclusion of each single observation.}

\footnote{For instance, Portes (1998) and Eichengreen and Portes (1987) suggest that generalized financial crises are of three sorts: banking crises, debt crises and foreign exchange crises. Radelet and Sachs (1998) distinguish between speculative attacks on the exchange rate, financial panics (or bank runs), collapses of asset price bubbles, moral hazard crises, and debt overhangs. Brière \textit{et al.} (2012) differentiate currency crises, sovereign debt crises, equity or bond crash crises, corporate bankruptcies or lost of confidence (Enron-type crises), and severe external events (9/11-type crises).}

\footnote{More broadly, the reference class problem may affect any empirical study regardless of its methodology. This applies in particular to model-free economic studies. For instance, data mining has limited predictive power when it comes to economic crises (Berg and Pattillo, 1999), likely because of its atheoretical backward-looking design. While the early-warning-system methodology proposed by Kaminsky, Lizondo and Reinhart (1998) seems to provide better results (Frankel and Saravelos, forthcoming), it still fails to incorporate the structural breaks that characterize crises (Candelon \textit{et al.}, 2012). Despite their undeniable explorative interest, model-free studies are of little help to address the reference class problem and build a meaningful classification of economic crises.}
In this paper, we show that current approaches are insufficient to address future crises. Since the reference-class problem is mostly ignored by economists, existing studies not only rely on biased samples but also, and most importantly, fail to imagine crises of new natures. As a result, the classifications found in the literature change through time and are inconsistent with each other. We contend that only a more conceptual approach on crises may overcome these limitations. In addition, we argue that indeterminacy should constitute a major building block of the economic theory of crises. In the current state of economic knowledge, the outburst of crises remains largely unpredictable. Admitting this fact instead of changing theory after each crisis will avoid the pervasive “this time is different” syndrome (Reinhart and Rogoff, 2011).

To illustrate this point, we put forward some common distortions stemming from ad hoc analyses of crises. In particular, we pinpoint the consequences of neglecting averted crises. Such “unborn” crises, often referred to as “peso problems,” are typically acknowledged for in specific contexts, but ignored when it comes to crises classification on a large scale. We therefore advocate for a forward-looking approach based on theoretical models general enough to generate multiple outcomes, some of which have not been observed yet. Such models likely represent the best currently available option to address economic crises and tackle the reference class problem.

2. The Reference Class Problem

The reference class problem formulated by Reichenbach (1949, p. 374) may be summarized as follows: “If we are asked to find the probability holding for an individual future event, we must first incorporate the case in a suitable reference class. An individual thing or event may be incorporated in many reference classes, from which different probabilities will result. This ambiguity has been called the problem of the reference class.”

In practice, establishing the probability of an economic crisis reduces to delineating reference classes to which such crises belong. Based on these classes, one could then try to evaluate the probabilities associated with each of them. The reference class problem epitomizes the fact that there are always many possible reference classes. As Eagle (2004, p. 393) puts it, “The obvious problem is that competing reference classes yield different probabilities, with no reference class standing out as the ‘correct’ one. Not only does the event seem to have no determinate unconditional probability, but there is no guide for the rational agent to assign one based on evidence, despite many attempts to provide one.”
Grouping crises together is an issue that some authors deal with by considering each crisis as unique. The “this-time-is-different” syndrome underlined by Reinhart and Rogoff (2011) emphasizes that economic analysts tend to consider each new crisis as the sole member of a new class. Like Taleb’s (2007) black swans, crises are rare and dramatic events, far away from repetitive events that probability statements are about, such as dice throws, roulette games, or even plane crashes. In the roulette, getting a, say, 17 is rather rare (probability: 1/37), but no more exceptional than any other outcome. It is a rare event, but not a dramatic one. On the contrary, plane accidents are dramatic events, but unfortunately frequent enough to allow for inference on their probability with an admissible confidence range (and subsequently make recommendations for lowering their occurrence).

Any event, however, is unique in time and place, be it a dice throw or a plane crash. Only abstraction enables to move from single cases to a repeated idealized event. For example, we ignore space-time modifications when speaking of “throwing ten times the same dice”. Doing so, we deliberately erase some “insignificant conditions” in order to gather similar events and search for general principles. “Categorical repetitivity of the world’s facts is a classificatory abstraction [through which] rational knowledge begins, [i.e.] when facts are reduced into classes as symbol of conceptual categories, with the search for laws which transcends the unique event.” (Scardovi, 1988, p. 59)

Single cases do not make science. But to which extent is categorization legitimate? At which point do we lose the specificity of the phenomenon we wish to study by putting it in a predefined class inappropriately? In games of chance, ignoring the “nitty-gritty causal details” (Millstein, 2003, p. 1321) of concrete situations (physical forces, etc.) in favour of symmetry considerations to elaborate scientific theories seems reasonable. Moreover, concerning roulette or plane crashes, everyone knows precisely what is at stake, no one is going to include “18” in the “getting a 17” event, no one is going to include car accidents in the plane crash statistics.

The same is not true for economic crises as each so-called crisis is typically associated to a single-event story (Bernal et al., 2010). Consequently, the issue goes far beyond “details” to be ignored in order to study crises. The question “How to define a crisis?” cannot be addressed without previously delineating what does matter, and what does not, in the
classification of crises. A good model for economic crises\(^5\) should then allow evaluating the *probability of crises*, with the aim of making policy recommendation for lowering their occurrence.

The economic literature offers various ways to establish the probability of crises in specific settings. In each case, the probability is derived from the classes to which the crisis belong (Canova, 1994). Formally, let the occurrence of a crisis be represented by the random variable \(X\), and let \(A, B, C\), etc. denote the different classes that have been previously identified. A model is intended to produce conditional probabilities, \(P(X|A)\), \(P(X|B)\), \(P(X|C)\). However, the probability of interest is the unconditional one, \(P(X)\), which represents the probability that a crisis will occur in the future, independently from the class it may belong to. Except if \(A, B, C\), etc. constitute a perfect partition of the set of possible crises (mutually exclusive and depicting any potential case), which may reveal impossible to reach in practice. If so, the determination of \(P(X)\) remains unsolved.

In a nutshell, the relative frequentist approach to crises occurrences is flawed. Such approach considers that the probability of event \(X\) is to be deduced from the relative frequency of \(X\) within a reference class. It implies a sort of “backward causation from future results to current chances” (Hàjek, 1997, p. 216). Such inference acts as if probability statements about \(X\) were counterfactually dependent on the future behavior of \(X\). But the probability of crises does not depend on whether there will actually be such crises in the future. For instance, crises linked to future technology by definition never occurred before the technology at stake was invented. Therefore, this approach is not the right concept for estimating the probability of a crisis.

As long as crises are not properly defined within unambiguous reference classes in which inference is possible, the “probability of a crisis” will remain pointless. In modern economic theory, observations are considered as outcomes of a generating probability distribution. However, associating probability distributions to single events is far from obvious. Philosophers do even quarrel about the existence of such probabilities (Settle, 1977; Milne, 1986; Levy, 2007; Bauman, 2005, 2008).

3. How do Economists Detect and Classify Crises?

\(^5\) see, e.g., Canova, 1994, and the references therein.
Classifications of economic crises come in many forms. Some authors track “abnormal” data, and declare that above/below a certain threshold there is a crisis. For instance, for Reinhart and Rogoff (2004) a free-falling episode is defined by a 12-month inflation threshold of 40% or more. For Reinhart and Rogoff (2011), an inflation crisis is an episode during which inflation exceed 20% per year, while hyperinflation means inflation above 500% per year.

Defining crises by means of thresholds and outliers has intuitive appeal. It gives the impression that crises share some features. This is however valid if and only if the underlying variables are driven by an objective and invariant probability distribution. One may wonder to which extent all so-called hyperinflations are indeed stemming from a unique distribution. Is there any reason to believe that the 1946 Hungarian hyperinflation was generated by the same underlying distribution as the one experienced by Zimbabwe in 2008? In some instances, the underlying structural relationship is obviously unstable. For example, Barkbu et al. (2011) stress the changes in both crises and IMF programs over time. Alternatively, one could argue that each relevant hyperinflation episode is the result of a conjunction of economic circumstances driven by country-specific factors. Some Latin-American countries have lived for long with two-digit inflation rates while European countries would view such a situation as a disaster.

Threshold-based definitions of crises bear the risk of being ad hoc. They lack theoretical grounds. For instance, data-driven thresholds for delineating hyper-inflationary episodes aren’t based on any sensible rule. Moreover, thresholds are arbitrary. Why would crises start at 20% inflation? Why not 19% or 21%? Although robustness checks can at least partially address this issue, threshold-based definitions of crises are both inelegant and unconvincing.

Most importantly, the threshold-based approach inevitably neglects crises averted just before their outbreak. Looking only at symptoms is insufficient to provide preventive cures. In medicine such an approach would have ruled out research aiming to understand why some patients get a disease and others remain immune. In the hyperinflation example, scrutinizing countries that managed to combat borderline-high inflation rate could be as insightful as restricting the analysis to those who fell in the trap.

---

6 Hyperinflation is just an illustration. Setting a threshold to define a crisis is probably the most common approach to economic crises.
A refinement to the threshold-based approach combines it with expert knowledge. In this line of thought, Schularick and Taylor (2012) use a two-step process to identify banking crises. First, they select “events during which a country’s banking sector experiences bank runs, sharp increases in default rates accompanied by large losses of capital that result in public intervention, bankruptcy, or forced merger of financial institutions” (p. 1038). Second, they send their selected dates to international experts for confirmation. Even though this refined approach has strong points, it still fails to provide an *ex ante* definition of crises. The assumption behind this approach is that expert knowledge is based on some obscure -- and likely incommunicable -- theoretical fundaments. Indeed, if theory was clear, it would allow any economist to apprehend crises with no need for external confirmation. *Ex post* expert knowledge is thus no quick fix to the reference-class problem.

From a more theoretical standpoint, Chang and Velasco (2001) distinguish three generations of models of crises: The first-generation model considers crises as an inescapable outcome of fiscal imbalances and fixed exchange rates. In the second-generation model, central banks may decide to abandon an exchange rate peg when its defense is too costly in terms of employment. In the third-generation model, the role of financial institutions and the domestic banking system is stressed. Classifying a single event in two different classes can bring polar consequences. In this respect, the discussion on the nature of the 1931 German crisis provides a good example. Should this crisis be analyzed with the first or the third generation model? According to Temin (2008), this difference matters not only for theoretical analysis but also for policy purposes. Indeed, if the German crisis of 1931 was just a currency crisis, the blame would rest on the government’s shoulders. Alternatively, if this crisis qualifies as a twin crisis, bankers would bear the responsibility. Thus, the classification of the 1931 German crisis affects both economic interpretation and policy recommendations. Another example is provided by Barkbu et al. (2011) who discuss the impact of grouping the Russian crisis of 1998 with the subsequent Latin American and Turkish crises. They conclude that “any taxonomy of crisis episodes is controversial.” (Barkbu *et al.*, 2011, p. 5).

Overall, issues related to class reference are amplified by the lack of consensus within the economics profession, even regarding the key determinants of hypothetical crises. As pointed out by Lo (2012), there is no agreement on the basic facts related to the latest crisis, let alone
on its starting date. Empirical evidence is fuzzy and interpretations as well as conclusions exhibit huge heterogeneity.

4. Lack of Imagination, Sample Bias, and Peso Problem

Even if all the economic crises were intrinsically similar in nature the current approaches in economics are biased and incomplete to cope with this issue. Indeed, investigations regarding crises fail to apprehend the whole universe which crises are drawn from. First, the typical backward-looking approach economists tend to adopt through econometric investigation lacks imagination about the emergence of new crises which would be the first one of their kind. Such new crises can result from new technologies or new ideologies. Within an ever evolving human society, the shapes of economic crises inevitably change. Fixed exchange-rates regimes entail other risks than floating ones. Internet banking allows current bank runs to remain unnoticed much longer than in the past. A sovereign debt crisis in a single-currency zone, like the Euro-zone, lies beyond previous experiences. Ignoring these realities stems for a lack of imagination.

In addition, *ex post* views on crises prevent researchers from factoring in events that were perceived as dramatic when they occurred but failed to leave a long-lasting footprint. In particular, the so-called “Peso problem” relates to *ex-post* implications of a disruptive event that was expected to happen, but actually did not\(^7\) (Veronesi, 2004). As Sill (2000, p.4) puts it: “Peso problems can arise when the possibility that some infrequent or unprecedented event may occur affects asset prices. The event must be difficult, perhaps even impossible, to accurately predict using economic history.”

In practice, this problem creates sample biases in econometric analyses of crises. Indeed, a Peso problem occurs when positive probabilities were initially assigned to ultimately *missing* events. In such a case, the actual observations used to calibrate or estimate models do not properly account for the real-time evidence, and predictions based on these models prove inaccurate. This is related to the philosophical ambiguity in the reference class problem. The

---

\(^7\) The name of this phenomenon is often attributed to Milton Friedman’s comments on the Mexican Peso market in the early 1970s (Sill, 2000). At that time, investors thought there was a probability that Peso would sharply devaluate, which was reflected in prices, but devaluation did not ultimately occur.
Peso problem underlines the fact that historical-data-based classes may be insufficient to cover any type of crisis that may occur. Deducing reliable probabilities from such classes is therefore inappropriate. As Peso problems are likely more frequent under dramatic circumstances where uncertainty about political interventions is high, historical data become unreliable in the context of crises. Therefore, classifying crises thanks to historical data can be misleading.

In addition, sample biases may result from deliberate choices made by researchers. For instance, many economists just discard wartime data. They implicitly assume that such data are not generated by the same economic mechanisms as “normal” (i.e. peace-time) data. Focusing on “normal data” however raises the issue of defining “normality.” The assumption that wartime data is “abnormal” from an historical standpoint may look ludicrous. As pointed out by Lerner (1954, p. 506), “economists have left whole areas of valuable statistical information unexamined because they have considered wartime figures “abnormal”’. Fifty years after this statement, most authors keep omitting wartime observations when dealing with long-term economic series. Such omission can result in severe sample biases, among which the survival bias.

The lack of imagination, the Peso problem, and the exclusion of “abnormal data” may yield inappropriate understanding of the nature of economic crises. Based on real-life situations, the remaining of this section shows how these three issues reinforce the reference class problem.

4.1. Lack of imagination

The 1918 repudiation by the Soviets of the debts issued by the Tsarist regime illustrates the incapacity to imagine new form of crises before they occur. This bond crisis likely represents the first historical example of ideologically motivated debt repudiations.

At the end of the 1890s France allowed Russia to issue a huge number of government bonds on the Paris Stock Exchange. Political reasons played a key role in this respect. As tensions with Germany escalated, the French government wanted to make sure that its Russian Ally would remain faithful. The Russian government, knowing the importance of external funding, heavily bribed French politicians (including the French finance Minister), the head of the Stock Exchange and the press (Raffalovitch, 1931). As a result, by 1914, Russian bonds were
traded on the Paris Bourse for an amount estimated to represent a staggering 4.5% of French wealth (Ukhov, 2003). When the Soviets managed to take power, they refused to recognize the debts issued by the Tsarist regime and decided, in February 1918, to repudiate all former Russian debts. Despite this extreme statement, bondholders kept hoping to be reimbursed. Landon-Lane and Oosterlinck (2006) show that more than two years after the repudiation some bonds were still traded above 50% of par.

French investors who had bought Russian bonds were certainly not irrational. When the repudiation occurred they had legitimate reasons to hope they would get reimbursed. Indeed at the time, there was no historical precedent in which a country had, solely for political reasons, repudiated a huge amount of debt without coming sooner or later to a settlement. Even in the dramatic hours following the French revolution, the legality of the Royal debt had not been questioned. The revolutionaries had defaulted on two-thirds of the debt but bondholders had nonetheless recovered part of their investment (Aftalion, 1996). Scholars who in 1917 worked on crisis could not possibly include the probability of an at-the-time-unknown form of crisis in their work: a Soviet revolution. The literature on sovereign debts considered three main reasons why countries would repay their debt: fear of loss of reputation (and thus loss of access to the capital market), fear of military intervention and fear of trade retaliations. The Russian case is exceptional in the sense that the Soviets did not fear any of these threats and thus the repudiation completely fell out of the range of considered outcomes.

The problems presented in the Russian case are in fact quite common. When the first commodity bubble centered on tulips imploded in the 17th century, observers were unable to understand that such a crisis could have occurred at all. They blamed wild speculation and irrationality. In the absence of a precedent in which commodities’ prices would have followed a bubble shape pattern, they were at a loss to include such a crisis in their expectations. The same may be said for the South Sea and Mississippi bubbles. In both cases, schemes to exchange government bonds for shares of companies ended up in a new form of crisis. Even the 1929 stock market crash may be viewed as a totally new form of crisis. World famous economists such as Irving Fisher completely failed to apprehend the crisis which would lead to a general dramatic drop of the prices of all the shares traded on the exchange.

Irving Fisher was found saying « Stock prices are low » in the New York Times on October 22, 1929, a few days before Black Thursday.
The current sovereign debt crisis in the Euro-zone provides another enlightening example. The risks related to the creation of a single-currency zone with no lender of last resort were hard to fathom. In fact, some scholars believed that the advent of the Euro zone had allowed countries to overcome their “original sin” problem, i.e., their incapacity to issue debt abroad in their own currency. Eichengreen et al. (2003) stressed the positive effects of the Euro on sovereign debt issuance. Their measure of original sin for Euro-land countries experienced a dramatic drop following the introduction of the Euro from (0.53 to 0.09). In contrast, when the crisis occurred, it appeared that countries in trouble were unable to print more Euros to reduce the debt burden.

4.2. Peso Problem

In today’s perspective, the reactions of the French bondholders following the Soviet debt repudiation may seem extremely incongruous. More than 80 years after the repudiation, we know that it took close to seventy years to reach a very limited agreement with Russia. In 1918 however, bondholders had legitimate reasons to hope for a reimbursement. Several events could indeed have dramatically changed the outcome for them. On the military side they could hope for a victory of the White Armies. On the political side, they could believe that the Soviets would agree to negotiate to be able to tap again international capital markets, to resume international trade or to be recognized as a legitimate government abroad. The creation of new countries such as Poland or the Baltic states on the ruins of the Russian empire meant that, following the practice in international law, part of the debts should be taken over by these countries. Eventually, some French political parties favored a partial bailout by French banks or by the French state. Landon-Lane and Oosterlinck (2006) detail each of these reasons to hope and show how logical it was for bondholders to keep hoping. In practice researchers working on basis of Russian bond prices only would, if they were not aware of all these issues, conclude that the crisis occurred much later than 1918. This is due to a Peso problem.

Peso problems also arise in various contexts where agents expected a crisis which eventually never occurred. Market prices then reflect pessimistic anticipations without any observable crisis associated. What would have happened if Cuba’s missile crisis or the “year 2K bug” had materialized? The range of such possible outcomes will never be explored, which inescapably hamper the scientific investigation of economic crises. Admittedly, this pitfall
affects not only economics but all sciences involving human reactions and expectations. But that should not make economists feel better about.

Several papers\(^9\) attempt to understand market anticipations regarding conflicts outcomes. In some cases, sharp market movements are not associated with events historians view as dramatic. According to Willard \textit{et al.} (1996) the most important break in the prices of greenbacks during the US civil war corresponds to the retreat of Confederate general Jubal Early from the suburbs of Washington, an episode largely downplayed by modern historians. Samples of crises build from \textit{ex post} historical appreciation thus likely omit crises perceived as such in real time. This is due to interpretations made with knowledge of later events. As pointed out by Frey and Waldenström (2007, p. 3), “this knowledge may bias the evaluation of the events, and may lead to “facts” being overlooked or overemphasized, as the case may be.”

\textbf{4.3. Abnormal data}

Economists often discard wartime data considering it “abnormal”. History shows however that the “abnormal” wartime can in some instances be the norm. During the period of Louis XIV’s reign when he ruled alone (1661-1715), France was at war more than 60\% of the time. For the Napoleonic period (1800-1815), the figure gets close to 80\%. Figures get even higher in the case of Philip II of Spain: During his reign (1556-1598), the empire was at peace for just one year (Drelichman and Voth, 2011). Furthermore, although wars usually take place in a well-delimited time frame, beginning and end dates are often blurred and economists fail to reach a consensus on these (Lo, 2012). Studies should thus be comprehensive enough to encompass all periods. Omitting “abnormal” data also prevents analyzing economic crises in their natural context.

Additionally, disregarding wartime data can trigger a survivorship bias. We illustrate this on the so-called “equity premium puzzle.” In their seminal paper, Mehra and Prescott (1985) argue that the size of the equity premium observed from US equities cannot be explained by standard general equilibrium models unless agents exhibit unrealistically high levels of risk aversion. This puzzle subsequently became a key issue in finance and macroeconomics

(Kocherlakota, 1996). A convincing solution to this puzzle is proposed by Jorion and Goetzmann (1999). Their argument draws on the survivorship bias associated to sample selection. They pinpoint that most empirical studies concentrate on uninterrupted markets only, and therefore observe equity returns much higher than if all markets were taken into account. To prove their point, Jorion and Goetzmann (1999) build a database comprising both interrupted and uninterrupted markets. They find that the equity premium is indeed significantly lower in interrupted markets. The deliberate omission of “abnormal” data may thus be sufficient to create a long-lasting “puzzle”.

We argue that the only way to overcome the three issues exemplified in this section is to start from theoretical models comprehensive enough to address the reference class problem. Such models should allow researchers to use all available data without distinctions. While sample biases and Peso problems are likely to affect any empirical study, they may be reasonably well circumvented when empirical analysis follows from a proper theoretical framework.

5. The Quest for a More Conceptual Approach

In natural sciences, explanatory power and realism are deeply related. The success of theories in physics relies on empirical confirmation, mainly through experimentation. In that context, the explanatory power is bound to the realism of the model: What the model explains is the reality under study. Because the models are empirically verifiable or falsifiable, their explanatory capacity merges with a descriptive and predictive capacity. The models concurrently explain and describe the world, as it is, and predict how it will be. Inescapably, models evolve with time, but this is the best researchers can do. As put by Batisty and Domotor (2008, p.169): “True, a model that works may still prove to be inadequate in some respects in the future. But accepting the mere possibility of such a future discovery does not in any way undermine our rationality as scientists in our commitment to the best-working model we currently have.”

Experimentation is sometimes unfeasible, like in the economics of crises. Then, the simultaneity of explanation and prediction is possible only if the phenomena under study are stable and repetitive, like the movements of planets. When astrophysicists model these movements, and make accurate predictions about them, they actually favour several characteristics that are already known, and obliterate others, so that mathematical models can
formally reproduce and organize them. In statistical physics, the kinetic theory of gases relies on numerous random particle movements. These are, in a sense, stable, because their future occurrence can, on a macroscopic scale, reasonably well be approximated by hypotheses made about their past behaviour.

From that perspective, prediction in the natural sciences is closer to “retrodiction” (Lambert, 2009), i.e., to the formal reproduction and organization of known phenomena. Higgs’ boson is a particle that was “mathematically discovered” before it was observed. But the underlying theory needed to be confirmed by observation before becoming part of the scientific corpus. Unfortunately, in economics there is currently no approach which would lead to the discovery of such thing as Higgs’ boson. So far, at best, one can expect some past stylized facts to corroborate current theories on crises, with no guarantee that the context will remain the same.

By focusing solely on past crises, even if one includes Peso problem episodes, researchers will miss new forms of crises occurring because of a change in the environment. Theoretical models have been put into question following the recent financial crisis. Obviously most models failed to predict this crisis. This failure is more likely due to a lack of imagination (getting beyond attempts to understand only past crises) than to a lack of intrinsic value of models. Noteworthy, and as pointed out by Franz (2011), failures to predict accurately crises is not a specificity of economics but is also quite frequent in natural sciences.

A nice, parsimonious, and consistent model is presently the best economists can produce about the emergence of crises.\(^{10}\) This could stem from a paradigm shift as advocated by Coats (1969). Rigorous models have the merit of helping clarify things and offering explanations consistent with pre-existing confirmed theories. Mathematical coherence is a valuable asset, especially when data are highly unreliable. Only in such a framework should economists try to come up with a definition and a classification of crises, and hopefully agree on them.

Nevertheless, predictions and policy-oriented recommendations derived from such models should be taken with caution. Even the most insightful model cannot rule out the high degree of uncertainty surrounding the occurrence of rare events. Like epidemiologists who have

\(^{10}\) Interestingly, Eisenhart (1989) develops a similar argument about theorizing case studies in management sciences.
announced the possibility of global pandemics, which in the end did not show off (mad cow disease, avian flu), economists find themselves in an uncomfortable situation when dealing with crises (Franz, 2011). Weighing the risk of panic against the accusation of being incapable of predicting future crises represents a challenging trade-off. In this respect, economists are in a worse situation than epidemiologists because scientific reputation is culturally more fragile for social than for natural scientists.

For natural scientists, realism is the most important issue. For social scientists, explanatory value is. Unification-type explanations are more suited to the social sciences than the deductive-nomological explanations of natural sciences (De Scheemaekere, 2009). Explanations indeed increase our understanding of the world by reducing the number of independent phenomena that we have to accept as ultimate. Hence, an economic theory has a significant explanatory capacity in so far as it describes the economic world with fewer independent phenomena – and thus makes this world, other things equal, more comprehensible (see also Friedman, 1974). This is especially relevant when data are unreliable.

It might be of great interest for a theory to account for once expected large changes that did not ultimately occur (Peso problem situations). Given the radical uncertainty underlying financial crises, the balance between the urge to predict (realism) and the need to understand (explanatory value) should tip in favour of the latter, because what matters eventually are the human decisions taken for lowering crises occurrences, not the (impossible) accurate prediction of the next crisis. Because there is no “retrodiction” in economics, the link between a theoretical model and its empirical verification rests on the soundness of a human judgement, not on the precision of a mechanical measurement.

Therefore a good model for crises should not first and foremost be able to predict the probability of occurrence of the next crisis. It should provide the tools to analyze the internal mechanisms of crises, without being contradictorily compelled to yield precise empirical predictions. This opens the door to models and theories that offer multiple possibilities and implications, without being considered as underspecified. Despite knowing that precise predictions are out of reach, economists are committed to come up with explanations for how to cure and prevent crises.
6. Are Multiple-Solution Models the Solution?

Randomness and probability may arise in science from two different perspectives: ignorance and intrinsic indeterminacy. On the one hand, if the world is deterministic and probability is due to the finiteness of human knowledge, then probabilities are typically human or subjective. This conception originates in Laplace’s (1840) conception on games of chance, which launched the mathematical probability theory (De Scheemaeckere and Szafarz, 2009). On the other hand, if probabilities belong to the world’s very nature, as in quantum mechanics, then they are objective and unavoidable even with absolute precision in measurement.

Concerning crises, the controversy may be stated as follows: Is the sudden and unexpected emergence of a crisis an intrinsic feature of our economies or a testimony to our inability to understand how the system works and properly predict its fluctuations? Gillies (2000) argues that objective probabilities do not apply to economics, because economic randomness only stems from sampling, not from the reality under study. In other words, economics is a deterministic science, like classical mechanics when opposed to quantum mechanics, the “epsilon” being there for the exclusive privilege of econometricians. From that common point of view, crises are not random. Only human failure to see them coming makes them look like random.

---

11 In the literature, the philosophers of probability agree that the various interpretations of probability can be divided into two broad groups, without agreeing on how these groups should be named. On the one hand, probability is called epistemic, epistemological, subjective, objective in the Platonic sense, or abstract. On the other hand, it is called objective, physical, empirical, scientific, random, statistical, or concrete (Gillies, 2000; Galavotti, 2005).

12 The typical example is provided by Young’s experiment, where a particle sent on a screen cut with two holes is found to have a fixed probability of going through one hole, whatever the number of experiments. This probability is objective in the sense that it is a feature of the particle’s state before meeting the screen, i.e., a feature of the world.

13 Gillies (2000) illustrates his point with an example: “At any given time, the proportion of families in England with budgets in a given range is a perfectly definite, though perhaps unknown, number. It is a frequency, not an objective probability. By taking a random sample we can use probability calculations to estimate this unknown number, but the objective probabilities in these calculations are introduced by the random sampling procedure and do not occur in the reality under study” (p. 193).
One could alternatively argue that economic dynamics are mostly unpredictable as suggested, in particular, by the financial theory of efficient markets.\textsuperscript{14} Scardovi (1988, p. 56) summarizes the argument: “This is not so much due to our intellectual limitation as to the intervention of accidental processes which are the very origin of variability and do not admit arcane predestinations, determinate necessities or immanent purposes.” Under that perspective, crises do not result from a predetermined, though hidden, reality. They are embedded in the underlying structure, which is random by nature. Hence, the “epsilons” are also part of pure theory. Furthermore, if economics is indeed stochastic by nature, then the level of indeterminism in crises occurrence is even more radical than in quantum mechanics, for we do not even know the range of possibilities, i.e., the probability distribution, associated with crisis-type events.\textsuperscript{15}

The trick here is human behaviour which is lying behind economic outcomes. Typical crash drivers, like self-fulfilling expectations and panics do not happen in physics.

In this line of though, indeterminacy becomes part of economic reality. Multiple-solution models suit well the representation of multifaceted reality. They bring distinct outcomes derived from the same economic relationships. Multiple solutions also naturally open the way to classification. Because the set of all solutions to a single model are well delineated, it becomes easy to partition this set into mutually exclusive sub-sets. Regarding crises, multiple-solution models may help distinguishing crises versus no-crisis situations. Moreover, among the crises, the classification could go one step further and introduce “crises of type A”, crises of type B”, etc. Models with multiple solutions may thus reveal fruitful in addressing the reference-class problem..

The previous section has demonstrated that statistical \textit{ex post} classifications often lack consistency. In contrast, using a single model that produces both crisis and no-crisis solutions would restore the supremacy of economic logic into the picture. A model encompassing all possible scenarios including dramatic ones, is not only more realistic, but also fitter to analyze potential consequences of different economic policies.

\textsuperscript{14} In a strongly efficient market, the asset price movements are unpredictable even when all relevant pieces of information are known by the economic agents (Jensen, 1978).

\textsuperscript{15} In quantum physics, the probability distribution is known from the physical features of the setting (the two holes in Young’s experiment for example), and can be tested empirically.
Considering multiple solutions as an attractive feature of economic models contradicts the vast majority of theoretical papers, which strive for obtaining a unique solution, often at the cost of artificial and unrealistic restrictions. These restrictions are typically based on stability assumptions. In contrast, we argue that stability restrictions harm the model specification because unstable solutions are key to address the reference-class problem. Instead of rejecting unstable solutions, one could interpret them in terms of crises. Imposing stability conditions is equivalent to ruling out crises beforehand.

Multiple solutions are nevertheless no quick fix to the reference-class problem. Two additional ingredients are needed. First, the model at stake should be carefully designed to endogenously capture the essence of different categories of crises, some of which have never materialized yet. Second, the parameters of the model should be identifiable from observed data series to allow for empirical estimation and tests. Luckily enough, the economist’s toolbox includes such models. As a matter of facts, most dynamic stochastic models admit multiple solutions.\footnote{Stochastic economic models vary according to the way agents are supposed to react to information. While rational expectations models constitute the dominant paradigm in economic theory, they are challenged by behavioral models built on the assumption that agents’ reactions are driven by some irrational beliefs often linked to known psychological biases. In principle, the reference class problem can be addressed in both frameworks. We will nevertheless illustrate the way crises can be addressed under the rationality assumption (Muth, 1961).} In particular, rational expectations models admit multiple solutions (Gray, 1984; Broze \textit{et al.}, 1985), some of which may display extreme statistical features. Moreover, Masson (1999) argues that multiple solutions are needed to properly address contagion, a typical feature of crisis periods.

The theory of speculative bubbles is a direct by-product of this mathematical result. In its simplest univariate form, the underlying model expresses the current price of a given asset as a linear combination of its future expectation and exogenous variables (see the Appendix for technical details). This model entails one so-called “fundamental solution,” interpreted as the price prevailing under normal circumstances, and infinity of “speculative bubbles,” which are associated with crises. In this simple case, there is only one no-crisis possibility, while any other outcome would belong to a single class of crises denominated as speculative bubbles. Some authors, like Salge (1997), have gone a bit further and identified different classes of bubbles.
7. Conclusion

To this day there is no consensus regarding the definition and classification of crises. This is reminiscent of the reference-class problem formulated by Reichenbach (1949). As exemplified in this paper, the reference-class problem has detrimental implications. In particular, attributing a given crisis to one or another class can lead to dramatically different conclusions, including for policy recommendation.

Typically, economic studies use an *ex post* threshold-based definition of crises. This implicitly assumes that all crises are but tail observations driven by the same underlying probability distribution as the “normal situation.” This contradicts the intuition that crises exhibit features that make the underlying model behave differently than during “normal” times. In particular, the threshold-based approach excludes speculative bubbles and self-fulfilling expectations, which are typically observed during dramatic situations.

From a theoretical standpoint, we pinpoint several limitations which add to the severity of the reference problem in the analysis of crises. First, focussing only on past crises and disregarding averted crises which were eventually averted is akin as building a medical science by *ex post* examining incurable patients only, and not those who either did not get sick or were successfully cured in due time. Second, economists lack imaginative models to apprehend new forms of crises. Rather, they tend to restrict their models to either “normal” and or “abnormal” past periods. This distinction ignores that reality is a succession of both normal and crisis periods. In particular, crises start in normal periods. Whereas in natural science, models can predict the existence of particles before they are empirically discovered (like the Brout-Englert-Higgs (BEH) boson) or diseases before their outbreak (like disease mutations), Existing economic models thus fail to consider potential innovations in crises.

This paper argues that multiple-solution models offer a promising avenue to address the reference-class problem plaguing the economic analysis of crises. Within those models, a given crisis would just be one possible outcome. Moving from non-crisis to crisis episodes and *vice versa* would be possible. Inevitably, multiple solutions yield indeterminacy. Such indeterminacy might however prove a force. Potential crises come from the theory itself, and not from *ex post* – and typically *ad hoc* – considerations. Models supposed to hold only under “normal conditions” and leaving the possibilities for crises aside are too narrowly conceived.
More generally, the use of indeterminacy in economic theory is debatable and likely not neutral. Economists have indeed their own objectives and rationality when producing knowledge. Davies and McGoey (2012) argue that ignorance is sometimes a convenient excuse for the absence of reaction to potential crises, and ultimately for the denial of responsibility. Moreover, models with multiple solutions never faithfully describe reality. Rather, they encompass it. They provide a large spectrum of possible outcomes, including some which will never be observed.

Admittedly, multiple-solution models constitute but one possible approach to produce an *ex ante* definition of crises. They nevertheless have the merit of showing that alternative ways to define crises while addressing the reference class problem exist. To seriously envisage such alternatives, economists should first abandon their suspicion against indeterminacy. Like in quantum physics, indeterminacy may indeed be a valuable source of conceptualization. In a sense, economic theory already benefits from conceptual building blocks pointing in that direction. Making an analogy with Kurosawa’s Rashomon, Lo (2012) suggests that economists may have to get used to the idea of living with “multiple truths”. Instead of fearing indeterminacy, economists should rather view it as a — somewhat unexpected — opportunity.
References


Appendix: Speculative Bubbles

By definition, speculative bubbles are solutions of a rational-expectation pricing model. Take for instance the following univariate model:

\[ p_t = \frac{1}{1+r} E[p_{t+1} | I_t] + x_t \]  

(1)

Where \( p_t \) is the price of an asset, \( E[p_{t+1} | I_t] \) is the rational expectation of the next-period price, \( r \) is the rate of return required for holding this asset, and \( x_t \) summarizes all exogenous factors affecting the determination of the current price.

Interestingly, model (1) admits an infinite number of solutions (for any given initial condition) although it entails no artificial indeterminacy in its basic specification. The existence of multiple solutions is due to the endogenous nature of rational expectations, which translates into the possibility for self-fulfilling expectations. The general solution of model (1) writes:

\[ p_t = p_t^f + (1+r)^t M_t \]  

(2)

Where \( p_t^f = \sum_{i=0}^{\infty} \frac{1}{(1+r)^i} E[x_{t+i} | I_t] \) is the so-called fundamental solution, and \( M_t \) is a martingale, i.e. a stochastic process such that: \( \forall i \geq 0: E[M_{t+i} | I_t] = M_t \). Because martingales may take infinitely many shapes, the general solution in (2) is compatible with an infinite number of price dynamics. Speculative bubbles then correspond to all solutions for which \( M_t \neq 0 \) (Flood and Garber, 1980; Blanchard and Watson, 1982).

Speculative bubbles offer an appealing way to define crises in a purely theoretical framework. Moreover, this approach applies to any model involving rational expectations, which is the case of most prevailing economic models. Strikingly, many authors tend to disregard this possibility by imposing a technical condition (namely, the so-called “transversality condition”) that rules out the possibility for speculative bubbles.\(^{17}\) This restriction is often justified by simplification motives only, although speculative bubbles are sensible outcomes of pricing models. Szafarz (2012) shows that bubbles arise in any financial market with speculators, regardless of the presence of fundamentalists.

\(^{17}\) The transversality condition is no necessary condition in a infinite-horizon framework (Szafarz, 2009)