WP/17/282

IMF Working Paper

A Model to Assess the Probabilities of Growth, Fiscal, and Financial Crises

by Suman S. Basu, Marcos Chamon, and Christopher Crowe

INTERNATIONAL MONETARY FUND

 WP/17/282

IMF Working Paper

Research Department

A Model to Assess the Probabilities of Growth, Fiscal, and Financial Crises*

Prepared by Suman S. Basu, Marcos Chamon, and Christopher Crowe

Authorized for distribution by Jonathan D. Ostry

December 2017

Abstract

This paper summarizes a suite of early warning models to assess the probabilities of growth, fiscal, and financial crises in advanced economies and emerging markets. We estimate separate signal-extraction models for each type of crisis and sample of countries, and we use our results to generate "histories of vulnerabilities" for countries, regions, and the world. For the global financial crisis, our models report that vulnerabilities in advanced economies were rooted in the bursting of leveraged bubbles, while vulnerabilities in emerging markets stemmed from lengthy booms in credit and asset prices combined with growing weaknesses in the corporate and external sectors.

JEL Classification Numbers: E37, F47, G01

Keywords: Early warning model, crisis probability, growth crisis, fiscal crisis, financial crisis, sudden stop

Authors' E-Mail Addresses: sbasu2@imf.org; mchamon@imf.org; ccrowe@capulaglobal.com

* For helpful discussions at various stages of this project, we thank Viral V. Acharya, R. Sean Craig, Atish R. Ghosh, Yuko Hashimoto, Jonathan D. Ostry, Mahvash S. Qureshi, Natalia Tamirisa, Charalambos Tsangarides, and participants at the 2014 IMF Conference on "Modeling Risks and Spillovers at the IMF" and at the 2016 Federal Reserve Board Round Table on "Financial Stability Monitoring and Early Warning Systems." For helpful discussions and excellent research assistance on the crisis probability models, we thank Eun Sun Jang, Hyeon Ji Lee, and Chifundo Moya, the latter of whom also helped prepare this working paper.

Contents

Appendices Table

I. INTRODUCTION

Crises are rare events, but when they occur, they generate large economic and social costs, they test the creativity and fortitude of policymakers, and most of all, they evoke the uncomfortable feeling that something should have been done to prevent them. Each wave of crises in the world is different, and sparks a new bout of soul-searching. Most recently, the global financial crisis of 2008–09 and the European sovereign debt crisis of 2010–12 established that crises are not just the preserve of emerging markets (EMs) and low-income economies (LICs), but can afflict a host of advanced economies (AEs) as well.[1]

In 2001, the IMF set up a Vulnerability Exercise (VE) to assess crisis vulnerability in EMs—an effort which was commended by the Independent Evaluation Office for having foreseen which EMs were most fragile at the time of the global financial crisis (IEO, 2011).[2] In the case of EMs, staff were worried specifically about external crises, so the first VE model, based on Berg and Patillo (1999), assessed the vulnerability of countries to currency crises, and the revised model that was introduced in 2007 assessed vulnerability to sudden stops (IMF, 2007). For an empirical model of crises in AEs, a broader view of crises needs to be taken, given the lack of sudden stops in these economies prior to the global financial crisis. This paper summarizes a suite of early warning models which were developed by staff in 2009 to assess the probabilities of growth, fiscal, and financial crises in AEs, and which were then extended in 2013 to assess the probabilities of these same types of crises in EMs.[3]

We wish to convey three main messages in this paper. Firstly, we describe how growth, fiscal, and financial crises differ from each other—plotting the dynamics for macro-financial aggregates such as the current account, real GDP, the fiscal balance, and private sector credit—and we document that the crises occur separately as well as together. Therefore, it is worth designing separate early warning models for each type of crisis.

---

[1] The macroeconomic phenomena that are grouped in common parlance as "crises" do actually vary quite significantly, both in terms of intensity and in terms of the sector of the economy that is affected. For emerging markets, the Latin American debt crisis of the 1980s, the Mexican crisis of 1994–95, and the Asian financial crisis of the late 1990s focused attention on banking and currency crises (e.g., Kaminsky and Reinhart, 1999) and sudden stops (following Dornbusch et al., 1995, and Calvo, 1998). For advanced economies, the recent crisis has sparked research on the heterogeneous growth effects of banking crises (e.g., IMF, 2009a, 2009b, and Romer and Romer, 2015) and on the fiscal crises induced by sudden stops (e.g., Merler and Pisani-Ferry, 2012).

[2] IEO (2011) states that for EMs, "the IMF gave consistent warnings on vulnerabilities related to: overheating, large current account deficits, credit booms, and unsustainable debt build-up," with the caveat that in surveillance documents, "still, in most countries, the overall messages were overly positive." The IEO's report criticized the IMF for not having extended the VE to AEs before the global financial crisis.

[3] The crisis probability models described in this paper are just one input into the VE for AEs, coming alongside a range of other sector-specific modules that are calculated using different methodologies. For EMs, our models are again just one input into the VE, coming alongside sector-specific scores that are all derived from an updated version of the (separate) EM sudden stop model. For LICs, there are separate early warning models for growth and financial distress. For a history of early warning models at the IMF up to 2012, see Chamon and Crowe (2013); for a comprehensive summary of the elements of the current VE, see Ahuja et al. (2017).

Secondly, the framework that we use is a signal-extraction model, building on the insightful contribution of Kaminsky et al. (1998), and our model turns out to be driven by indicators of whether leveraged bubbles are building and bursting, and by measures that reflect risk sharing, market adjustment, and policy response in the aftermath of shocks. Given our focus on a wide variety of crises, we use a large range of indicators outside the external sector, which was the central focus of the literature on sudden stops (this body of work identified reserve coverage, the current account deficit, external debt, and indicators of exchange rate overvaluation as the main contributors to vulnerability). The in-sample fit of our models is slightly better than the models evaluated by Berg et al. (2005): on average, the sum of the percentages of missed crises and false alarms in our models is 0.37.[4]

Our third message in this paper is that our models can be used to generate a "history of vulnerabilities" for each country, which in turn can be aggregated to produce a global history with reasonable properties.[5] The global average and spread of crisis probabilities appears to increase prior to each wave of crises, and for the global financial crisis, our models report that vulnerabilities were high in AEs because of bursting leveraged bubbles, while EM vulnerabilities stemmed from corporate and external sector weakness coupled with ongoing credit and asset price booms. At the individual country level, although the upper range of predicted probabilities goes up to 0.50, probabilities rarely exceed 0.15, and when they do, they often feature a sharp upward spike followed by a speedy decline. In AEs, this decline typically comes after a bubble has stopped deflating and external imbalances have adjusted, while in EMs, it comes alongside a recovery in corporate, public, and external sector health.

The remainder of this paper is organized as follows. Section II rationalizes our model design in relation to the existing literature and outlines our contributions. Section III covers the signal-extraction procedure that is applied in both the AE and EM models. Sections IV and V delve into the crisis dynamics, estimation details, and aggregated results for AEs and EMs respectively. Section VI outlines some of the nuances and limitations that we feel important to bear in mind when applying our models in practice. Finally, section VII concludes.

---

[4] The percentage of missed crises is equal to the percentage of crisis observations that our model fails to flag as vulnerable, while the percentage of false alarms is equal to the percentage of non-crisis observations that our model incorrectly flags as vulnerable. Section II discusses the relationship between in-sample and out-of-sample fit, while sections III, IV and V describe the construction of the crisis flags underlying the above statistic.

[5] We take several steps in this note to preserve the confidentiality of the VE results: (i) we do not present any country-specific results; (ii) we only present back-cast results from our model based on our latest estimation dataset, not the ratings that were actually produced and/or assigned in the VE rounds; (iii) we present global/ regional averages only up to 2014 (so the recent post-taper-tantrum period is excluded); and (iv) to prevent the identification of vulnerability ratings for large countries, none of the aggregate results are GDP-weighted.

II. Model Design

In this section, we explain three key decisions in designing our model—namely, which crises to focus on, which estimation method to use, and whether to focus on flagging vulnerabilities to crises or on the timing of crises—in the context of the early warning literature and past IMF experience in VE analysis.[6]

Type of crisis

We can categorize early warning models into two "generations." The first generation was inspired by the EM crises of the 1980s and 1990s. Most of these papers assessed the risk of speculative attacks, involving some subset of exchange rate depreciation, loss of reserves, and increase in the domestic interest rate (e.g., Frankel and Rose, 1996; Eichengreen et al., 1995; Kaminsky et al., 1998; Glick and Hutchison, 1999), or of a sudden stop in capital inflows (e.g., Chamon et al., 2007).[7] Other papers focused on banking distress (e.g., Borio and Lowe, 2002; Eichengreen and Rose, 2004).

The second generation of models was inspired by the global financial crisis and attempted to assess vulnerabilities to a large range of crises in both AEs and EMs. External crises were still an important focus (e.g., Obstfeld et al., 2009; Catão and Milesi-Ferretti, 2014), but so were collapses in GDP and equity returns (e.g., Rose and Spiegel, 2010, 2011; Blanchard et al., 2010; Berkmen et al., 2012; Frankel and Saravelos, 2012). Christofides et al. (2016) observed that the different dimensions of the global crisis—i.e., distress in banking, balance of payments, exchange rates, and growth—could not all be explained by a single variable, but that each kind of crisis is associated with different risk factors.

This paper contributes to the second generation of early warning models, exploring crises other than sudden stops. Given the financial nature of the trigger for the global crisis, we are interested in banking crises. In addition, given the large output collapses across the world, we also model directly a sudden slowdown in GDP growth from its long-term trend. Finally, we need to empirically assess the risk of a fiscal crisis, but given the lack of sovereign debt distress events in AEs prior to 2009, we develop a new crisis measure that captures episodes of sudden fiscal tightening. We conjecture that such events are motivated by large negative reassessments of the long-term fiscal outlook. We estimate separate models for each type of crisis; therefore, our approach is consistent with the insights of Christofides et al. (2016).

---

[6] This section draws on Chamon and Crowe's (2013) discussion of the evolution of IMF VE models, as well as on additional literature produced since that work was published.

[7] Some studies focused on a joint incidence of different types of crises, such as Kaminsky and Reinhart (1999), which examined simultaneous currency and banking crises, and Ghosh and Ghosh (2003), which restricted attention to those currency crises that were also associated with a large decline in real GDP growth.

Our choice of variables go hand in hand with our crisis selection. Recent history suggests that the risk of growth, fiscal, and financial crises may be larger when there has been a sudden and large downward reassessment of the sustainability of previous booms in credit and asset prices, whether those booms were financed domestically or externally. We therefore include two types of variables: medium-term and near-term. Our medium-term variables capture sustained booms in credit and asset prices, which may be liable to reverse if they turn out to reflect bubbles of some kind; while our near-term variables encompass sudden reductions in asset price growth (which are a sign that a bubble may be about to burst) as well as indicators of the resilience of the economy to bursting bubbles and other kinds of shocks.[8] In terms of post-shock resilience, we seek to capture a range of variables summarizing the scope for risk-sharing (e.g., the reliance on debt versus equity), private-sector cushioning (e.g., corporate liquidity buffers), and policy measures (e.g., the space for fiscal and monetary stimulus).

Estimation method

While several early warning models have used regression analysis (either linear or limited-dependent-variable estimations), we choose instead to use a modified version of the non-parametric signal-extraction approach pioneered by Kaminsky et al. (1998). These authors calculate a threshold value for each variable, such that observations on one side of the threshold are flagged as risky while those on the other side are flagged as safe. The threshold is calculated to maximize the signal of the flag relative to its noise; this means minimizing some combination of false alarms (the percentage of the risky flags that is not followed by a crisis) and missed crises (the percentage of the safe flags that is followed by a crisis).[9]

Our choice is motivated by two reasons. Firstly, both Berg and Patillo (1999) and Berg et al. (2005) find that out of the models they evaluate, only the model by Kaminsky et al. (1998) has reasonable out-of-sample properties.[10] Secondly, the signal-extraction approach is more practical given the limitations imposed on us: (i) each variable's threshold is calculated separately, so if a variable is missing in one year, it just makes that variable's threshold less precisely calculated, but it does not mean that other variables that are available in that year cannot be used; (ii) since each threshold is typically located in the middle of the values of that variable, it is robust to poor data quality in outliers; (iii) we can aggregate the signals over a large number of variables, while a similar number of variables in a regression analysis may generate substantial multi-collinearity;

[8] The importance of credit growth in the early warning context is affirmed by a range of studies (e.g., Borio and Lowe, 2002; Schularick and Taylor, 2012; Gourinchas and Obstfeld, 2012; Babecký et al., 2012; Alessi and Detken, 2017). Drehmann and Juselius (2014) design a new measure of the debt servicing burden that, together with the credit to GDP ratio, performs well in flagging vulnerabilities to banking crises.

[9] We calculate each variable's threshold separately. An alternative approach is a binary classification tree, where each variable's threshold at any point in the tree depends on whether other variables have been flagged as above or below their thresholds (e.g., Ghosh and Ghosh, 2003; Chamon et al., 2007; Alessi and Detken, 2017).

[10] Berg and Patillo (1999) estimate the models of Kaminsky et al. (1998), Frankel and Rose (1996) and Sachs et al. (1996) using data through 1996 and assess their ability to forecast events in 1997. Berg et al. (2005) compare out-of-sample properties for their own model, the Kaminsky et al. (1998) model, and two private sector models.

and (iv) the thresholds and their calculation are relatively easy to interpret and to communicate to policymakers.

Berg and Patillo (1999) and Berg et al. (2005) convincingly established that in-sample fit may be a poor representation of out-of-sample performance, and we have attempted in our model design to address this finding. Firstly, we limit the use of interaction terms only to the terms for household leverage and asset prices, because allowing unrestricted complexity in variable design may provide spuriously high in-sample fit. Secondly, we use the same set of variables for each type of crisis while allowing the thresholds and weights to differ, instead of excessively tailoring the model, because the latter would generate the risk that we capture past crises well but not new ones which may strike in different ways. Thirdly, and most importantly, we have to grapple with how to use the knowledge about AE crises that we gained from the global financial crisis, but without overfitting our AE model such that it only explains the global financial crisis and nothing else (given the much lower frequency of AE crises in other time periods). Our compromise has been to use variables such as household leverage that were revealed to be important by the global financial crisis, but to restrict the estimation procedure for AEs to the pre-2008 period.

Vulnerability versus timing

At the time of the 2007 revision of the VE, the IMF established a distinction between crisis vulnerabilities and crisis triggers (IMF, 2007), which was re-affirmed in Ghosh et al. (2009), IMF (2010), and Ahuja et al. (2017). Vulnerabilities are factors that increase the economy's exposure to shocks (e.g., excessive growth in credit, balance sheet mismatches) while crisis triggers are those shocks (e.g., political news, bankruptcies in key sectors, contagion from abroad) which interact with vulnerabilities to generate crises.

Conceptually, we task our model with assessing vulnerabilities. When we calculate a crisis probability, we provide a translation from an index measure of vulnerability (achieved by aggregating crisis flags over individual variables) into a metric that is easier to understand, but this metric is unconditional on current events and based purely on crisis frequency over our historical sample. Crisis triggers, on the other hand, are extremely difficult to predict using an early warning model because they are highly non-stationary across periods, and so we make no attempt to predict their timing. Judgment from outside the model is necessary to convert our probability into a probability conditional on the latest news about shocks.

In any case, on a practical dimension, communication with economic policymakers is better confined to vulnerabilities than crisis triggers and timing. Firstly, once vulnerabilities are flagged, policy actions can be undertaken to reduce them before triggers materialize. Secondly, if vulnerabilities are persistently high, and the way that this information is conveyed to policymakers is that a crisis is imminent, then there may be "signal fatigue" after a while, and policymakers may ignore the model's results simply because no triggers have occurred, leaving the economy more vulnerable to a trigger once it eventually strikes.[11]

In this paper, we use our model to construct a "history of vulnerabilities" for each country, for each region, and for the world, and we view these histories as being of independent interest—we do not just check whether crises occurred in a way that matches the histories. Since crisis occurrence requires both vulnerabilities and triggers, we can use the information from the crises that we have observed, in order to assess whether the world was vulnerable to crises in other periods as well, even when (luckily) no triggers materialized.

III. Estimation Procedure

There is substantial heterogeneity in shocks, adjustment mechanisms, and data availability between AEs and EMs. Recognizing this, our approach is to select a different set of variables for each group of countries and then to conduct three separate estimations, one for each type of crisis, for each group of countries. Within each country grouping, the set of explanatory variables is mostly identical across the three estimations, but the different estimations assign different thresholds and weights to the variables.

Since we choose variables from across all sectors of the economy, each of our estimated crisis models allow for inter-sectoral spillovers in reduced form (e.g., *financial* sector variables may be estimated to increase the *fiscal* crisis probability if in the historical sample, financial sector distress has caused the realization of public sector contingent liabilities and necessitated a sudden fiscal tightening). And for all our estimations, we use annual data—therefore, we relate variable values in each year to crisis incidence in the following year.

In this section, we outline the crisis definitions, loss function, threshold estimation, aggregation procedure, and probability conversion that is applied in both the AE and EM models. In the following two sections, we describe the country sample, crisis dynamics, variable selection, and aggregated results for AEs and EMs respectively.

Growth, fiscal, and financial crises are defined as follows:

- Growth crises are defined as occurring when the change in growth relative to the previous five-year average growth rate lies in the lower 5th-percentile tail.
- Fiscal crises are defined as occurring when there is a year-on-year increase in the fiscal balance of 2.5 percent of GDP or higher, starting from a balance of -2.5 percent of GDP or lower.

[11] Financial sector institutions may be more interested in the precise timing of crises, since that timing affects the optimal timing of transactions. In that case, higher-frequency models combined with judgment based on market-specific knowledge may be necessary to complement an early warning model such as ours.

- Financial crises are defined as systemic banking crises and/or currency crises, following Laeven and Valencia (2012). These authors define a systemic banking crisis as an episode with "significant signs of financial distress in the banking system (as indicated by significant bank runs, losses in the banking system, and/or bank liquidations)" and "significant banking policy intervention measures in response to significant losses in the banking system." They define a currency crisis as "a nominal depreciation of the currency vis-à-vis the U.S. dollar of at least 30 percent that is also at least 10 percentage points higher than the rate of depreciation in the year before."[12]

Following Berg et al. (2005), our loss function is the unweighted sum of the percentages of missed crises and false alarms. Each observation takes the form of a country-year pair plus associated readings across a set of variables. The percentage of "missed crises" is defined to be the percentage of crisis observations that our model fails to flag as vulnerable, while the percentage of "false alarms" is defined to be the percentage of non-crisis observations that our model incorrectly flags as vulnerable. Given the rarity of crisis events, this definition of the loss function imposes the specific value judgment that missing a crisis is substantially more costly than issuing a false alarm (e.g., if crises are 5 percent of the sample, then missing one crisis is as costly as issuing 19 false alarms).

For each explanatory variable, the value of the threshold is chosen so as to minimize the above loss function.[13] We illustrate the threshold selection decision graphically in figure 1, by plotting the cumulative distribution functions (CDFs) of crisis and non-crisis observations against the value of the variable in the previous year. We assume that a higher value of the variable is associated with a higher risk of crisis, so the crisis CDF lies to the right of the non-crisis CDF. For a given threshold, we classify all country-year observations with variable values above the threshold as vulnerable, and all observations with variable values below the threshold as not vulnerable. This means that the height of the crisis CDF at the threshold represents the percentage of missed crises, while the height of the non-crisis CDF at the threshold represents the complement of the percentage of false alarms. The optimizing threshold is therefore the one where the vertical gap between the CDFs is maximized.

Each country-year observation is assigned a separate flag for each variable (1 if the variable value is on the vulnerable side of the above threshold, and 0 otherwise), and those flags are then aggregated across variables to yield a single composite vulnerability index score for that country in that year. We weight each variable using its signal-to-noise ratio:

$(1-z)/z$, where z = sum of the percentages of missed crises and false alarms.[14]

---

[12] For more details on the definitions of systemic banking and currency crises, together with a discussion of the necessary judgments involved in diagnosing systemic banking crises, see Laeven and Valencia (2008, 2012).

[13] Our threshold estimation differs from Kaminsky et al. (1998): our thresholds are a common value of the variable across all countries, while their approach is to first convert each value of the variable observed within each country into a country-specific percentile, and then to find a common percentile across all countries. In section VI, we relate the appropriateness of each approach to the degree of heterogeneity within the sample.

[14] Notice that minimizing the loss function amounts to maximizing each variable's signal-to-noise ratio.

The weights are normalized so that they sum to unity, and the weight of missing variables is reallocated to the other variables in proportion to their weights. To avoid over-weighting variables with similar information, variables are grouped when the correlations between their 0/1 flags are above 0.4. The total group weight is determined by the signal-to-noise ratio of the most informative variable in the group, and each variable's relative weight is assigned in line with its relative signal-to-noise ratio.

The composite score is converted into a crisis probability using a non-parametric mapping.[15] Our procedure is outlined in figure 2. We first plot the history of crisis realizations (taking the value of 0 if there was no crisis, or 1 if there was a crisis) against the composite scores assigned by the model to each country-year observation. Then from those discrete values, we generate a continuous line representing the average of the 0/1 crisis realizations around each value of the composite score, by using a modified version of the HP filter (modified to apply over the score value, not over time). Finally, we obtain our desired mapping from composites to probabilities by imposing a monotonicity restriction on the continuous line. Our mapping may have flat portions, within which the score may vary without altering the probability.

IV. Advanced Economies

Our sample of AEs covers 33 countries from 1985 onwards.[16] As described in section II, our estimation period ends in 2007, in order to avoid the overfitting problem that would come with including the global financial crisis period within the estimation window. Figure 3 and table 1 document crisis incidence in the AE sample, and underline that the global crisis had a substantial impact: following a wave of financial crises in 2008, there was a wave of growth crises in 2008–09 and a wave of fiscal crises in 2009–12. The average frequency of crises in the sample rose from 0.03 in 1985–2007 to 0.17 during 2008–10.[17]

It is useful to have a separate model for each type of crisis because the different crises do not always occur together and because they appear to capture conceptually different phenomena. Figure 4 shows that while growth, fiscal, and financial crises have sometimes occurred together, they have also often struck on their own (out of the three crisis types, financial crises seem to have the most overlap, especially with growth crises). Figure 5 documents that the different crisis types are associated with different dynamics for macro-financial aggregates. Typical growth and financial crises appear to be cushioned by an expanded fiscal deficit, which is quite different to our definition of a fiscal crisis. Financial crises appear to be followed by worsening growth performance in subsequent years, while growth crises are followed by a quick recovery in activity.

[15] See Kaminsky (1999) for an alternative non-parametric approach with a similar motivation.

[16] Appendix Table A1 contains the list of countries in the AE sample.

[17] Financial crises are defined only up to 2010 by Laeven and Valencia (2012). Accordingly, for consistency in our comparisons between different types of crises, we restrict such comparisons to the period 1985–2010.

For our AE estimations, we use the set of variables shown in table 2. Taking advantage of good data availability in AEs, we use a range of variables that were revealed to be important by the global financial crisis. In the category of medium-term variables, we capture sustained booms in credit and asset prices as well as relative growth in the construction and financial sectors. In the category of near-term variables, we include sudden reductions in asset price growth (i.e., negative readings for the “acceleration” variables), indicators of the dependence of household balance sheets on asset price developments (i.e., the interaction terms featuring the ratio of household liabilities to GDP), and sector-level indicators of overheating and of resilience to bursting bubbles and other kinds of shocks. Every crisis model uses all the variables in the table, except for the primary gap and public debt variables (which only the fiscal crisis model uses).

The in-sample fit of our models varies across the three types of crisis, and on average is better than the models evaluated by Berg et al. (2005).[18] Our measure of in-sample fit comes from conducting a threshold estimation over the composite vulnerability index score that each model produces, and assigning flags according to that threshold (table 3). The financial crisis model performs best, and the growth crisis model worst, within the estimation period 1985–2007, and the average sum of the percentages of missed crises and false alarms is 0.30.

Of course, for practical purposes, out-of-sample fit is a better indicator of model performance than in-sample fit. But we do not have a good out-of-sample test, and testing our models on the 2008–2010 period is not an adequate substitute: firstly, our models were only developed after the global financial crisis, so they contain conceptual information relevant to that crisis; and secondly, our objective was to develop a model of general AE crises, not just the (rather unusual) global crisis. Nevertheless, comparing the in-sample fit of our models with the fit during the global financial crisis yields interesting economic lessons. While the fiscal crisis model’s fit remains the same as the in-sample fit, the growth crisis model performs worse, and the financial crisis model becomes completely uninformative.[19] This result suggests that while fiscal crises, and to a lesser extent growth crises, occurred during 2008–10 in a manner similar to the pre-crisis period, financial crises did not. Financial crises in that period may have been determined by variables that are not captured by our model—e.g., a pattern of banking-sector linkages between each country and troubled banks in the United States.[20]

[18] The lowest in-sample sum of errors recorded by Berg et al. (2005) was 0.58, while the lowest out-of-sample sum of errors was 0.63. Bear in mind however that the models evaluated by Berg et al. (2005) are not strictly comparable to ours: they are based on monthly data, on a different type of crisis, and on a different time period.

[19] If the composite score is completely uninformative, the crisis and non-crisis CDFs lie on top of one another, and the sum of errors is 1. The financial crisis sum of errors is 0.96, very close to 1, during 2008–2010.

[20] This observation is consistent with the post-crisis literature on banking-sector and trade-sector contagion effects. Cerutti’s (2013) model on banking-sector linkages has been incorporated into the VE, combining the crisis probabilities from our models with cross-border banking data from the Bank of International Settlements. Recent work on the linkages between the banking and shadow-banking sectors includes Abad et al. (2017).

Our models can be back-cast to generate a "history of vulnerabilities" for each country, taking the form of a time series of each crisis probability and a decomposition of each sector's contribution to that probability (figure 6).[21] Each year's probability is based on the data from the previous year. For the remainder of this section, we first show that the global and regional aggregates produced by our model are reasonable, and then we document the dynamics of crisis probabilities at the individual country level.

Figure 7 shows unweighted averages of global and regional crisis probabilities, as well as the ranges and interquartile ranges of the country-level probabilities, for the period 1990–2014.[22] At the global and regional levels, the mean probability of each type of crisis is elevated in the early 1990s, then increases in the run-up to the global financial crisis, and finally jumps again in 2012 (with the fiscal crisis probability rising especially strongly at that date). The dispersion—especially the upward spread above the median—of the country-level crisis probabilities appears to increase before and during the global crisis, as vulnerabilities accumulate asymmetrically across the sample of countries.[23]

We decompose each sector's contribution to the global crisis probability in figure 8. The relative sectoral contributions vary for the different types of crises, and evolve over time as different kinds of imbalances build up over different decades and as more data becomes available.[24] From our model's perspective, average vulnerabilities were high in the global crisis year of 2009 primarily because of household balance sheet effects (i.e., the bursting of asset price bubbles to which households were exposed because of their own leverage). Other important sources of vulnerability that were roughly co-equal in importance were the real, financial, and external sectors. The public sector became especially important from 2010 onwards, driving an increase in the fiscal crisis probability.

In figure 9, we zoom in on the vulnerabilities in those countries in Europe which received external assistance (from the IMF, EFSF/ESM, and/or ECB SMP) during the global crisis and the subsequent European sovereign debt crisis. Crisis probabilities jumped to a higher value for these countries than they did globally. While the global story about household balance sheet

[21] The crisis probability is decomposed into sectoral contributions by assuming that the percentage contribution of each sector to the probability is equal to the percentage contribution of variables in that sector to the composite score. The latter percentage is easy to calculate, since the composite score is the weighted sum of crisis flags across all sectors.

[22] These graphs are designed to describe our model outputs fairly well in a manner that preserves confidentiality of the VE results. For many reasons, they are not the best measure of overall global and regional risks (for a list of the constraints imposed by confidentiality, see footnote 2; for a more detailed discussion of the appropriate way to think about global and regional risks, including factors outside our model, see section VI).

[23] The median of the financial crisis probabilities never changes because a significant fraction of countries are clustered around a flat portion of the crisis probability mapping, at a probability of around 0.08. However, the mean and dispersion measures move similarly to the other types of crises.

[24] The sectoral contributions on the graph encompass information about the weight of each sector in the estimation, the frequency with which the thresholds in that sector were violated, the mapping of the threshold violation into a crisis probability, and of course, whether data on that sector is available for that year.

vulnerabilities holds true for these countries as well, the public and external sectors were more important for these countries than for the other countries in the sample.

Next, we turn to the dynamics of the individual country-level crisis probabilities that underlie the global and regional picture, using our model outputs over 1985–2017. To facilitate the analysis, we categorize the probabilities into low, medium, and high ratings (L, M, and H) using the cutoffs that are currently used for AEs in the VE.[25] Table 4 documents the Markov transition matrices for the three types of crises, and figure 10 uses the data in those transition matrices to calculate: (i) the long-run probabilities of countries being in each category; and (ii) the long-run expected duration of each category, conditional on starting there.[26] This analysis illustrates what we have ourselves seen from handling the crisis probabilities in each VE round: although the upper range on the probabilities goes up to 0.50 (figure 7), the probabilities rarely exceed 0.15, and when they do, they often feature a sharp upward spike followed by a speedy decline.

How do AEs get into and leave high crisis probability ratings? Figure 11 shows that over time, household balance sheet vulnerabilities have become more important among those countries who are diagnosed with high ratings, both because of secular trends on asset prices and balance sheet leverage, and because of the improved availability of these data. Figure 12 documents that those countries who have exited a high rating for growth and financial crisis have done so as a result of an improvement in household balance sheets (typically because an asset price bubble has stopped deflating) and a reduction in external imbalances. Exiting a high rating on fiscal crises tends to be associated instead with a reduction in public sector vulnerabilities.

V. Emerging Markets

Our sample of EMs covers 53 countries from 1994 onwards.[27] As documented in figure 13 and table 5, crisis incidence in the EM sample reached new heights in the global financial crisis period, but there did exist earlier waves of crises as well, particularly in the mid- and late 1990s. Accordingly, the issue of overfitting the global financial crisis is less salient, and our estimation period covers the full period 1994–2010. The average frequency of growth, fiscal, and financial crises rises from 0.04 during 1994–2007 to 0.07 during 2008–10.[28]

---

[25] Our categories are defined as follows: L = below 0.05; M = from 0.05 to 0.15; and H = above 0.15.

[26] We use transition matrices instead of just calculating the probabilities and expected durations from our sample because our sample may not have a long enough time period to calculate long-run averages accurately, especially for the rarer H category. However, we note that the matrix analysis relies on the Markov assumption.

[27] Appendix Table A2 contains the list of countries in the EM sample.

[28] We compare our EM models to sudden stops because early warning models for sudden stops in EMs are more common in the literature; such a model forms the backbone of the IMF's VE for EMs (see Ahuja et al., 2017, for more details). Following Chamon et al. (2007), we define sudden stops as episodes of significant declines in private net capital inflows, and we identify them using a combination of quantitative measures followed by judgment by IMF desk economists to rule out spurious events. The quantitative assessment requires any of the following to be satisfied: (i) private net capital inflows at least 1.5 standard deviations below their mean and at least 0.75 standard deviations lower than the previous year; (ii) private net capital inflows at least 1.5 standard deviations lower than the previous year and at least 0.75 standard deviations lower than two years before;

Because of the focus of the earlier EM literature on sudden stops, we next compare the incidence and dynamics of each type of crisis not just to each other as we did in the previous section, but also to sudden stops. However, all of our subsequent vulnerability estimations are conducted for growth, fiscal, and financial crises only.

Although there is a significant overlap of growth, fiscal, and financial crises with sudden stops, each type of crisis occurs separately and appears to capture conceptually different phenomena. Therefore, it is useful to have a separate model for each type of crisis in EMs, using a model distinct from the sudden stops model. Figure 14 documents that each type of crisis strikes on its own and, moreover, that the set of crises classified as sudden stops may actually contain rather heterogeneous events: some sudden stops are associated with growth collapses, others with sudden fiscal tightenings, and still others with financial crises. Only rarely do all types of crises occur together. Figure 15 plots the macro-financial dynamics around each type of crisis. Financial crises are most similar to sudden stops, but the current account adjustment is larger in the latter. Fiscal crises are unusual in being associated with only a small impact on the current account, and while a fiscal crisis involves a tightening of the fiscal balance, all other crises are cushioned by an expanded fiscal deficit. Sudden stops and financial crises are followed by worsening growth performance in subsequent years, while growth crises are followed by a quick recovery in activity.

For our EM estimations, we use the set of variables shown in table 6. Relative to the AE estimations, data availability is generally poorer, so we have fewer indicators related to household balance sheets and the construction and financial sectors. On the other hand, given our priors about the greater importance of the external sector for EM crises, we include measures of building and bursting bubbles in foreign investor sentiment (proxied by the medium-term appreciation in the real effective exchange rate, REER, and by sudden declines in the rate of appreciation), of the foreign appetite for domestic sovereign debt (i.e., the EMBI spread), and of the dependence of the economy on oil imports or exports. Like for AEs, every crisis model uses all the variables in the table, except for the primary gap and public debt variables (which only the fiscal crisis model uses).

The in-sample fit of our models has an average sum of errors of 0.44, which is slightly better than the models in Berg et al. (2005), but worse than the AE models in section IV (table 7). The fiscal crisis model performs best, and the financial crisis model worst, over the estimation period 1994–2010. The models perform better during the global crisis period, with the average sum of errors declining from 0.46 during 1994–2007 to 0.38 during 2008–2010.

What if we amend the estimation period for the EM models to end in 2007, as we did for the AE models? Then we could conduct an out-of-sample test for the models over the global crisis

(iii) private net capital flows at least 0.75 standard deviations lower than the previous year and at least 1.5 standard deviations lower than two years before; (iv) private net capital flows as a share of GDP at least 3 percentage points lower than the previous year and at least 2 percentage points lower than two years before.

period 2008–2010. Of course, it would not be a perfect test: the EM models for these types of crises were developed in 2013, with more information available to us than would have been available on the eve of the global crisis in 2007. Nevertheless, even an imperfect test may be informative, and so we present the goodness of fit measures for the truncated EM models in table 8. It appears that the change in the sample period does not alter very much the in-sample fit of the models. The average sum of errors over the out-of-sample period, 2008–2010, is 0.65. This is worse than in table 7, but still reflects a commendable model performance, and is in line with Berg et al.'s (2005) calculated out-of-sample performance for the Kaminsky et al. (1998) model. Unlike in the AE case, none of the crisis models become uninformative during the global crisis. For the remainder of this section, we return to reporting results for the EM models estimated over the full sample period 1994–2010.

We can again use our models to back-cast a "history of vulnerabilities" for each country (figure 16). Figure 17 shows the unweighted averages of global and regional probabilities, as well as the ranges and interquartile ranges of the country-level probabilities, for the period 1994–2014, subject to the same caveats that applied for the AE graphs in figure 7. At the global and regional levels, the mean probability of each type of crisis is elevated throughout the late 1990s and then increases with the global financial crisis. The dispersion—especially the upward spread above the median—of the country-level crisis probabilities appears to be highest at these times as well. Since crises tend to affect countries in the highest tail of the vulnerability distribution, this upward spread is very important.

Figure 18 decomposes each sector's contribution to the global crisis probability. The real sector (capturing corporate sector health and inflation) accounts for a declining source of vulnerabilities over time, perhaps owing to institutional improvements such as market liberalization and inflation targeting. Household balance sheets account for a much lower share of overall vulnerabilities than they did in the AE sample, partly because of poorer data availability for EMs. From our model's perspective, average vulnerabilities were high in the global crisis year of 2009 primarily because of roughly equal contributions from the real sector, external sector, and the medium-term variables capturing the credit and asset price booms that were ongoing at the time. Public sector vulnerabilities persisted after the crisis, causing the fiscal crisis probability to remain elevated.

In figure 19, we zoom in on the vulnerabilities in emerging Asia during the Asian financial crisis period. Growth crisis probabilities increased steadily from 1994 to 1999 on the back of a deteriorating corporate sector and inflation performance. Fiscal crisis probabilities were already high in 1995 because of external imbalances, then reached a low in 1998 before rising again as a result of weakening real, financial, and public sector health. Financial crisis probabilities grew from 1994 to reach a peak in 1997, the year of numerous speculative attacks on the currencies of East Asian countries. Our model views financial vulnerabilities to have arisen from excessive medium-term growth in credit and asset prices combined with external imbalances, and the model also picks up signs of a slowdown in asset price growth in 1996 which may have presaged the eventual crash.

Finally, we turn to the dynamics of the individual country-level crisis probabilities, using our model outputs over 1994–2017. We again categorize the probabilities into low, medium, and high ratings (L, M, and H).[29] Figure 17, table 9 and figure 20 establish that although the upper range on the probabilities goes up to 0.40, the probability and duration of countries at the high probability rating (above 0.15) are rare and fleeting, just like for AEs.

How do EMs get into and leave high crisis probability ratings? Figure 21 shows that over time, medium-term vulnerabilities related to credit and asset price booms have become more important among those countries who are diagnosed with high ratings. The improved availability of household balance sheet data is also visible. On the other hand, the prevalence of real sector vulnerabilities in these countries has declined. Figure 22 documents that those countries who have exited a high crisis probability rating have done so as a result of improvements in the real sector (particularly for financial crisis ratings but also for growth crisis ratings) and in the public and external sectors (for fiscal crisis ratings).

VI. LIMITATIONS

In this section, we outline the key nuances and limitations that we feel important to bear in mind when applying our models, based on our practical experience over several VE rounds.

Regime shifts

Any estimation procedure implicitly assumes that all countries within the sample follow a similar underlying stochastic process, or "regime," governing the conversion of explanatory variables into overall crisis vulnerabilities. We separate AEs and EMs for estimation purposes because they clearly follow different regimes as a result of differences in economic and political institutions. However, our models do not capture the possibility of differences of regime *within* each estimation sample.

On the economic front, there are two dimensions of possible heterogeneity in the regime. Firstly, there may be differences in market liberalization and institutional development between the countries in each sample (in particular, the EM sample covers a very diverse set of countries). The extent of these differences determines whether the estimated thresholds for each variable should be common across countries or country-specific.

Both approaches have drawbacks. Consider a non-crisis country whose public debt level has been slowly declining over a long period, but nevertheless remains high relative to other countries. A model like ours, with common thresholds, would identify the debt as indeed generating a vulnerability, but one that is possibly mitigated by the country simultaneously

[29] The cutoffs for EMs in the VE are: L = below 0.03; M = from 0.03 to 0.15; and H = above 0.15. The reason for the lower L/M cutoff relative to the AEs is technical: one of the probability mappings has a flat portion which is just above or below 0.05, and we do not wish our VE ratings to be excessively sensitive to the cutoff.

having safe levels of other variables. By contrast, a model with country-specific thresholds, such as Kaminsky et al. (1998), would be less likely to flag the debt as a vulnerability, and more likely to judge that this specific country, perhaps unlike others in the sample, can simply sustain such a debt level without experiencing a crisis. The country-specific approach is appropriate when regimes differ between countries, while our approach works better when some similarity in regime can be assumed, because it allows for the sources of vulnerability to be identified (in this example, the fiscal sector), which is particularly important from a policy perspective if the exercise is meant to inform efforts to mitigate those vulnerabilities.

Secondly, economic institutions may evolve within countries over time. During our estimation period, many EMs have liberalized their markets and developed new institutions for fiscal policy, monetary policy, and the exchange rate regime. Within the AE sample, the creation of the Eurozone has changed the scope for monetary policy and external adjustment for many European countries, while also integrating their financial sectors. In addition, some countries began as EMs and then evolved into AEs during the course of the sample period.[30]

To be clear, changes in institutions that are fully captured in the explanatory variables of our estimations do not pose a problem. For example, in EMs: the abandonment of fixed exchange rate regimes would have reduced measured overvaluations (in terms of the deviation of the current account from the value calculated by the IMF's External Balance Assessment, or EBA); the issuance of sovereign debt in domestic instead of foreign currency would have prevented external debt to GDP ratios shooting up after depreciations; and the adoption of inflation targeting would have reduced inflation rates. Some of these changes may indeed have been responsible for our model finding no upward trend in crisis probabilities in EMs despite a significant secular increase in capital flows and financial interconnectedness. In AEs, the higher capital flows following the introduction of the euro would have been captured by our models as well.

However, our approach does face a problem if changes in institutions are not fully captured by our explanatory variables. For example, consider an EM that has switched to exchange rate flexibility, domestic instead of foreign currency denomination of debt, and an independent central bank. In such a country, a large stock market decline that would have previously caused a run on the banks and on sovereign debt may no longer do so, because investors trust that an exchange rate depreciation can easily restore the solvency of the financial and public sectors without generating a debt overhang or an increase in long-term inflation expectations. For another example, consider an AE that adopted the euro in 1999. At a conceptual level, the loss of monetary policy and external adjustment capacity perhaps immediately increased the probability that a housing market shock would make the country susceptible to a growth collapse, a sudden fiscal tightening, and bank runs. However, our models are not designed to capture this kind of an

[30] This observation raises the question as to whether some portion of these countries' histories should be included in the EM estimation and other portions in the AE estimation.

increase in vulnerabilities (which would only be captured later, once there is an observed deterioration in the indicators related to overheating and leveraged bubbles).

Political regime shifts are also important. During turbulent eras such as in the aftermath of large slowdowns in activity, every election may come with a large risk of abrupt changes in policy. This means that the frequency of crisis triggers becomes elevated, and in addition, some of the cushioning mechanisms for economic activity may be removed—for example, bank bailouts may become politically toxic, or an exchange rate depreciation may be ruled out of bounds owing to inter-country tensions. As a result, the same economic vulnerabilities may become associated with a higher crisis probability.

Global risk outlook

The global averages presented in sections IV and V are good representations of the output of our model, but they are not necessarily the best measure of global risks. Partly this is because in order to preserve the confidentiality of country-specific results, in this paper we do not GDP-weight either the average probability or the upper tail of the probability distribution (this tail is where crises are most likely to be concentrated).

More fundamentally, our models are necessarily constrained in what they can say about global risks because they are based only on data from the individual country level (they were designed to assess individual country vulnerabilities, and from the perspective of individual countries, global developments fall into the category of external crisis triggers rather than domestic vulnerabilities). In order to build a picture of global vulnerabilities, we need additional information on the health of the key international financial and goods markets on which global economic activity depends, and on the space available for global policy adjustment. Some examples of useful indicators would be macro evidence on the building and bursting of global credit booms,[31] sectoral data on whether bottlenecks are building in important financial and goods markets, and assessments of the correlation and transmission of various kinds of shocks across countries.[32]

Data and estimation

The reliability of our crisis probability output ultimately rests on the appropriateness of the set of selected variables, the quality of the data, and the robustness of the estimation. Here we outline some of the issues that we have encountered with each.

---

[31] Recent papers on early warning models have incorporated measures of global growth and inflation (e.g., Babecký et al., 2012) and global liquidity (e.g., Alessi and Detken, 2017). These papers complement an emerging literature on the global financial cycle, following Rey (2015, 2016).

[32] As mentioned in section IV, our model is not able to capture the spread of financial crises in AEs during 2008–10, perhaps in part because it lacks information on the precise banking-sector linkages as they stood on the eve of the global crisis. Cerutti (2013) was instrumental in incorporating such evidence, and his model has been combined with our model's crisis probabilities to form the "banking contagion" module of the VE. The VE also combines our crisis probabilities with international trade linkages to form a "trade contagion" module.

Variable selection raises two distinct issues. Firstly, the appropriate number of variables to include is unclear. On the one hand, including too few variables may be undesirable, because we need to make sure that our model robustly captures all the possible ways in which a fundamental imbalance (in financial or goods markets) can manifest itself, and not just the specific manner in which it did manifest itself during a specific crisis wave in the recent past. For example, it may be easy to derive a high signal-to-noise ratio for a specific crisis wave using just stock price growth, but in the next wave, financial imbalances may manifest themselves through housing price growth instead. On the other hand, as described in section II, we have not allowed unrestricted complexity in the design of new variables, because that can also lead to spuriously high in-sample fit. As an extra check, we have designed the aggregation procedure to down-weight correlated variables.

Secondly, we choose variables for which relatively plentiful and reliable cross-country data is available over a lengthy period of time, but this comes with drawbacks. There is a good argument for mostly using data that is available over long time periods: choosing variables which have only become available more recently runs the risk of overfitting, as the model becomes better at predicting the most recent crises in the sample at the detriment of predicting crises more generally. However, this approach rules out the use of some recent data that has been associated with crisis incidence, e.g., sovereign CDS spreads.

Next, turning to data quality issues, we note that while a threshold estimation approach mitigates some problems, others remain. As described in section II, the threshold calculation is robust to poor data quality in outliers. However, each country's crisis probability becomes quite sensitive to any data measurement errors that may push the recorded variable value just above or just below the calculated threshold. In practice, this sensitivity has been particularly salient for balance sheet and survey data, when the variable in question is slowly revised over time as more observations enter our database. The initial reading for the variable may be just below the relevant threshold, while later in the year, it becomes revised to just above.

Finally, we turn to the robustness of the estimation process itself. One problematic issue that we have noticed is that for some variables, the calculated threshold value is liable to jump up or down from one VE round to the next. This phenomenon occurs when the distance between the crisis and non-crisis CDFs is near its maximum for two rather different values of the variable (figure 23 shows an example). While our aggregation process does account for the signal-to-noise ratio of each variable, it does not really account for this kind of uncertainty in the threshold itself, and the impact thereof on the reliability of the 0/1 crisis flags (the same value of the variable may be classified as vulnerable to crisis in one round and not vulnerable in the next). We can think of two approaches to deal with this issue. Firstly, we can use our judgment to simply select the threshold value that accords better with the existing literature or with the desired ratio of missed crises to false alarms. Secondly, we can generate bootstrapped estimates of the thresholds. This approach also produces estimates for the standard deviations of the threshold values, which could in the future be taken into account in the aggregation process.

Another issue is that while threshold estimation can be conducted without technical difficulty in an environment where a regression analysis might run into the multi-collinearity problem, the problem of correlated variables does not actually disappear. Running a variable-by-variable analysis creates a problem of omitted variable bias within each threshold calculation, which artificially inflates the weight of variables which happen to be correlated with other variables that were truly responsible for crises in the past. While our down-weighting of correlated variables does help to partially account for this issue, there is a deeper problem if the true risk factors are not included in the model—whether because they are forgotten or because they are not readily measurable.

VII. CONCLUSION

As we emerge from the long shadow of the global financial crisis, it is critical that domestic and international policy institutions, aided by the research of external academics, seek to diagnose in advance the signs of impending future crises, and to thereby catalyze policy actions that may prevent crises from materializing. In this paper, we have outlined the suite of early warning models that we have developed at the IMF since 2009 to assess the vulnerabilities of both AEs and EMs to growth, fiscal, and financial crises. Our crisis probabilities are comparable across countries and over time, and are used alongside a range of other modules in the IMF's Vulnerability Exercise.[33]

Our three main messages are as follows. Firstly, it is useful to have a separate model for each type of crisis, because the different crises do not always occur together and because they appear to capture conceptually different phenomena. Secondly, we use a signal-extraction approach to derive our crisis probabilities, using a set of explanatory variables covering indicators of whether leveraged bubbles are building and bursting, as well as indicators of risk sharing, market adjustment, and policy response in the aftermath of shocks. Thirdly, our models can be back-cast to generate a "history of vulnerabilities" for each country, which in turn can be aggregated to produce a global history with reasonable properties. The global average and spread of crisis probabilities appears to increase prior to each wave of crises, including the recent global crisis. At the individual country level, probabilities mostly stay at low or medium levels and rarely enter the high category, but when they do, they often feature a sharp upward spike followed by a speedy decline. At both the global and individual country levels, our early warning models can help decompose which sectors of the economy have historically been responsible for large increases in vulnerabilities, and which sectors have generally needed to improve before crisis probabilities can decline.

[33] To facilitate the use of our model by country desks and the communication of our results to country authorities, we provide in every VE round a spreadsheet that documents the time series for each country's crisis probabilities and the reasons that the probabilities changed since the previous round.

References

Abad, J., M. D'Errico, N. Killeen, V. Luz, T. Peltonen, R. Portes, and T. Urbano, 2017, "Mapping the Interconnectedness between E.U. Banks and Shadow Banking Entities," NBER Working Paper 23280 (Cambridge MA: NBER).

Ahuja, A., K. Wiseman, and M. Syed, 2017, "Assessing Country Risk—Selected Approaches," IMF Technical Notes and Manuals 17/08 (Washington: International Monetary Fund).

Alessi, L., and C. Detken, 2017, "Identifying Excessive Credit Growth and Leverage," *Journal of Financial Stability*, forthcoming.

Babecký, J., T. Havránek, J. Matějů, M. Rusnák, K. Šmídková, and B. Vašíček, 2012, "Banking, Debt, and Currency Crises: Early Warning Indicators for Developed Countries," European Central Bank Working Paper 1485 (Frankfurt, Germany: ECB).

Berg, A., E. Borensztein and C. Pattillo, 2005, "Assessing Early Warning Systems: How Have They Worked in Practice?" *IMF Staff Papers*, 52(3): 462–502.

Berg, A., and C. Pattillo, 1999, "Are Currency Crises Predictable? A Test," *IMF Staff Papers*, 46(2): 107–138.

Berkmen, S. P., G. Gelos, R. Rennhack, and J. P. Walsh, 2012, "The Global Financial Crisis: Explaining Cross-Country Differences in the Output Impact," *Journal of International Money and Finance*, 31(1), 42–59.

Blanchard, O. J., M. Das, and H. Faruqee, 2010, "The Initial Impact of the Crisis on Emerging Market Countries," *Brookings Papers on Economic Activity*, 2010(1): 263–307.

Borio, C., and P. Lowe, 2002, "Asset Prices, Financial and Monetary Stability: Exploring the Nexus," BIS Working Paper 114 (Basel, Switzerland: Bank for International Settlements).

Calvo, G. A., 1998, "Capital Flows and Capital Market Crises: The Simple Economics of Sudden Stops," *Journal of Applied Economics*, 1(1), 35–54.

Catão, L. A. V., and G. M. Milesi-Ferretti, 2014, "External Liabilities and Crises," *Journal of International Economics*, 94(1):18–32.

Cerutti, E., 2013, "Banks' Foreign Credit Exposures and Borrowers' Rollover Risks: Measurement, Evolution and Determinants," IMF Working Paper 13/09 (Washington: International Monetary Fund).

Chamon, M., and C. Crowe, 2013, "Chapter 34 - Predictive Indicators of Financial Crises," in *The Evidence and Impact of Financial Globalization*, G. Caprio, T. Beck, S. Claessens and S. L. Schmukler (eds.), San Diego CA: Academic Press.

Chamon, M., P. Manasse, and A. Prati, 2007, "Can We Predict the Next Capital Account Crisis?" *IMF Staff Papers*, 54(2): 270–305.

Christofides, C., T. S. Eicher, and C. Papageorgiou, 2016, "Did Established Early Warning Signals Predict the 2008 Crisis?" *European Economic Review*, 81: 103–114.

Dornbusch, R., I. Goldfajn, and R. O. Valdes, 1995, "Currency Crises and Collapses," *Brookings Papers on Economic Activity*, 1995(2): 219–293.

Drehmann, M., and M. Juselius, 2014, "Evaluating Early Warning Indicators of Banking Crises: Satisfying Policy Requirements," *International Journal of Forecasting*, 30(3): 759–780.

Eichengreen, B., and A. K. Rose, 2004, "Staying Afloat When the Wind Shifts: External Factors and Emerging-Market Banking Crises," in *Money, Capital Mobility, and Trade: Essays in Honor of Robert A. Mundell*, G. Calvo, M. Obstfeld, and R. Dornbusch (eds.), Cambridge MA: The MIT Press.

Eichengreen, B., A. K. Rose, and C. Wyplosz, 1995, "Exchange Market Mayhem: The Antecedents and Aftermath of Speculative Attacks," *Economic Policy*, 10(21), 249–312.

European Central Bank, 2013, "Press Release: Details on Securities Holdings Acquired Under the Securities Markets Programme," February 21 (Frankfurt, Germany: ECB).

Frankel, J. A., and A. K. Rose, 1996, "Currency Crashes in Emerging Markets: An Empirical Treatment," *Journal of International Economics*, 41(3-4), 351–366.

Frankel, J. A., and G. Saravelos, 2012, "Can Leading Indicators Assess Country Vulnerability? Evidence from the 2008–09 Global Financial Crisis," *Journal of International Economics*, 87(2): 216–231.

Ghosh, A. R., J. D. Ostry, and N. Tamirisa, 2009, "Anticipating the Next Crisis—What Can Early Warning Systems be Expected to Deliver?" *Finance and Development*, 46(3): 35–37.

Ghosh, S. R., and A. R. Ghosh, 2003, "Structural Vulnerabilities and Currency Crises," *IMF Staff Papers*, 50(3): 481–506.

Glick, R., and M. Hutchison, 1999, "Banking and Currency Crises: How Common Are Twins?" Pacific Basin Working Paper Series PB99-07 (San Francisco CA: Federal Reserve Board of San Francisco).

Gourinchas, P.-O., and M. Obstfeld, 2012, "Stories of the Twentieth Century for the Twenty-First," *American Economic Journal: Macroeconomics*, 4(1): 226–265.

Independent Evaluation Office of the International Monetary Fund, 2011, "IMF Performance in the Run-Up to the Financial and Economic Crisis: IMF Surveillance in 2004–07." (Washington: International Monetary Fund).

International Monetary Fund, 2007, "Assessing Underlying Vulnerabilities and Crisis Risks in Emerging Market Countries—A New Approach," SM/07/328, September 17 (Washington: International Monetary Fund).

International Monetary Fund, 2009a, "From Recession to Recovery: How Soon and How Strong?" Chapter 3 of *World Economic Outlook: Crisis and Recovery* (April): 97–132 (Washington: International Monetary Fund).

International Monetary Fund, 2009b, "What's the Damage? Medium-Term Output Dynamics after Financial Crises," Chapter 4 of *World Economic Outlook: Sustaining the Recovery* (October): 121–151 (Washington: International Monetary Fund).

International Monetary Fund, 2010, "The IMF-FSB Early Warning Exercise: Design and Methodological Toolkit," IMF Policy Paper (Washington: International Monetary Fund).

Kaminsky, G. L., 1999, "Currency and Banking Crises: The Early Warnings of Distress," IMF Working Paper 99/178 (Washington: International Monetary Fund).

Kaminsky, G. L., S. Lizondo, and C. M. Reinhart, 1998, "Leading Indicators of Currency Crises," *IMF Staff Papers,* 45(1): 1–48.

Kaminsky, G. L., and C. M. Reinhart, 1999, "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems," *The American Economic Review,* 89(3): 473–500.

Laeven, L., and F. Valencia, 2008, "Systemic Banking Crises: A New Database," IMF Working Paper 08/224 (Washington: International Monetary Fund).

Laeven, L., and F. Valencia, 2012, "Systemic Banking Crises Database: An Update," IMF Working Paper 12/163 (Washington: International Monetary Fund).

Lane, P. R., and G. M. Milesi-Ferretti, 2007, "The External Wealth of Nations Mark II: Revised and Extended Estimates of Foreign Assets and Liabilities, 1970–2004," *Journal of International Economics,* 73(2): 223–250.

Merler, S., and J. Pisani-Ferry, 2012, "Sudden Stops in the Euro Area," Bruegel Policy Contribution 2012/06 (Brussels, Belgium: Bruegel).

Obstfeld, M., J. C. Shambaugh, and A. M. Taylor, 2009, "Financial Instability, Reserves, and Central Bank Swap Lines in the Panic of 2008," *The American Economic Review: Papers and Proceedings,* 99(2): 480–486.

Rey, H., 2015, "Dilemma not Trilemma: The Global Financial Cycle and Monetary Policy Independence," NBER Working Paper 21162 (Cambridge, MA: NBER).

Rey, H., 2016, "International Channels of Transmission of Monetary Policy and the Mundellian Trilemma," *IMF Economic Review,* 64(1): 6–35.

Romer, C. D., and D. H. Romer, 2015, “New Evidence on the Impact of Financial Crises in Advanced Countries,” NBER Working Paper 21021 (Cambridge MA: NBER).

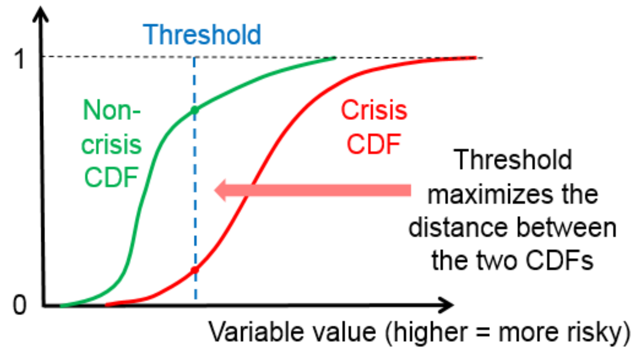
Rose, A. K., and M. M. Spiegel, 2010, “Cross-Country Causes and Consequences of the 2008 Crisis: International Linkages and American Exposure,” *Pacific Economic Review*, 15(3): 340–363.

Rose, A. K., and M. M. Spiegel, 2011, “Cross-Country Causes and Consequences of the Crisis: An Update,” *European Economic Review*, 55(3): 309–324.

Sachs, J., A. Tornell, and A. Velasco, 1996, “Financial Crises in Emerging Markets: The Lessons from 1995,” *Brookings Papers on Economic Activity*, 1996(1): 147–199.

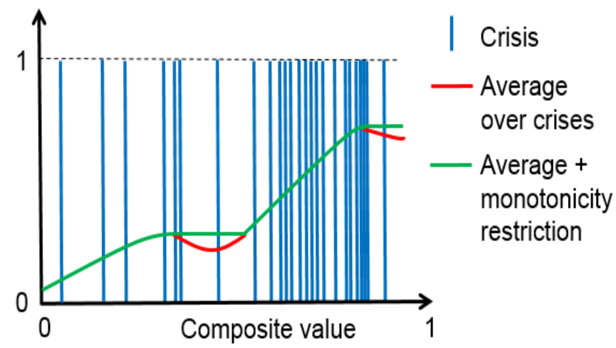
Schularick, M., and A. M. Taylor, 2012, “Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008,” *The American Economic Review*, 102(2): 1029–1061.

Figure 1. Threshold Estimation



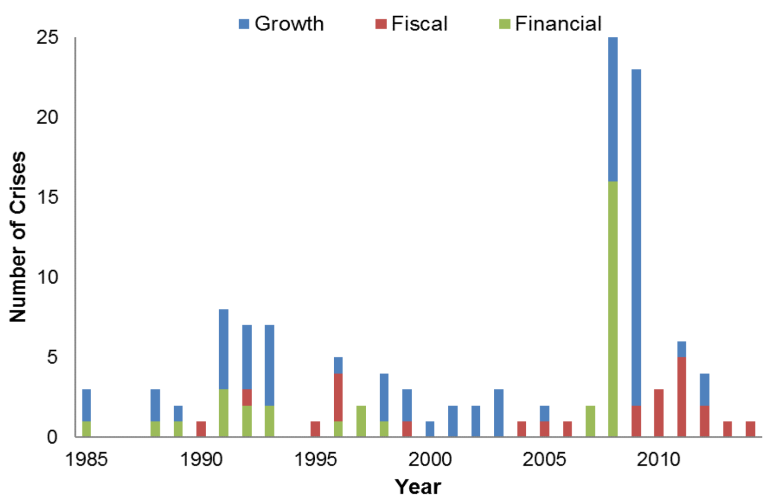
Source: Authors' illustration.

Figure 2. Probability Mapping



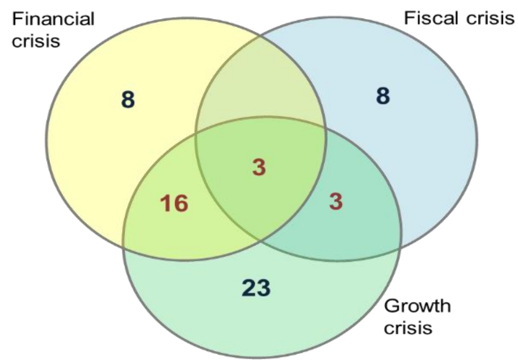
Source: Authors' illustration.

Figure 3. Crisis Incidence in AEs



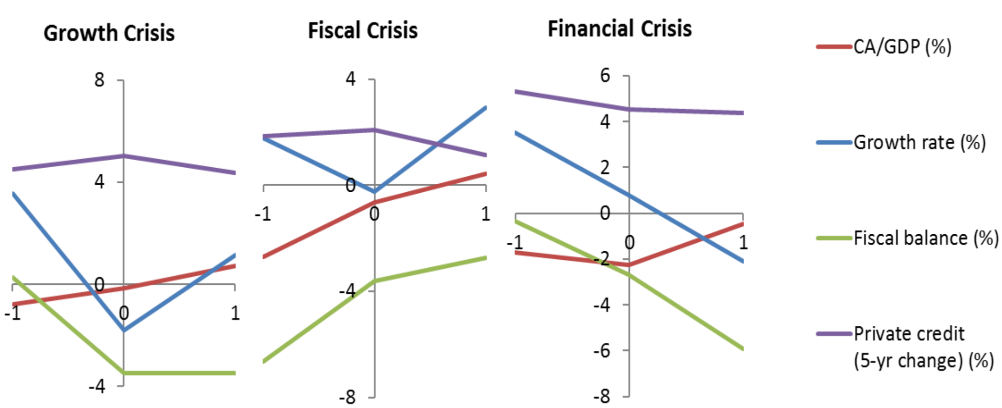
Source: Laeven and Valencia (2012) and authors' calculations.
Note: Data on financial crisis incidence is available only up to 2010.

Figure 4. Overlap of AE Crises, 1985–2010



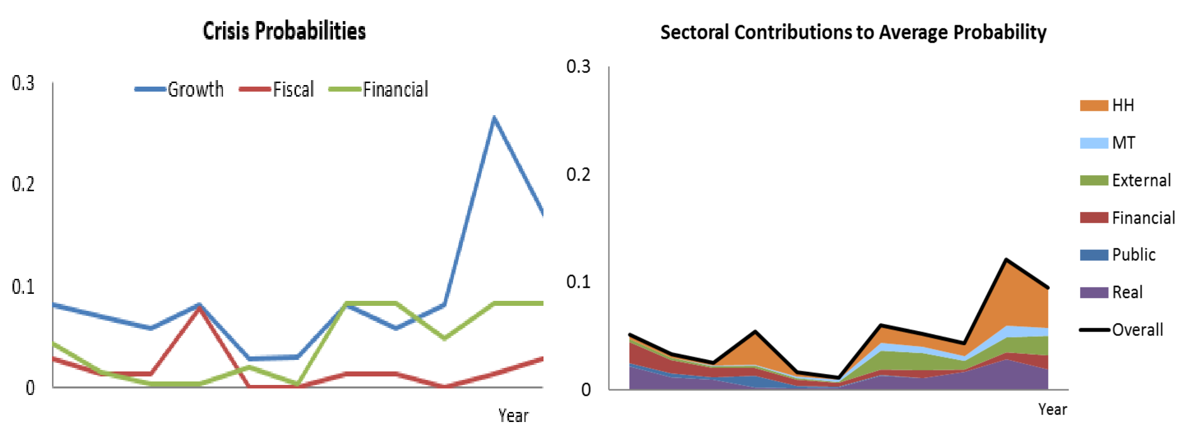
Source: Authors' calculations.
Note: Crises in the same or adjacent years are regarded as coincident.

Figure 5. Macroeconomic Dynamics Around AE Crises, 1985–2010



Source: Authors' calculations.
Note: Date 0 is crisis year; Dates -1 and 1 are years before and after.

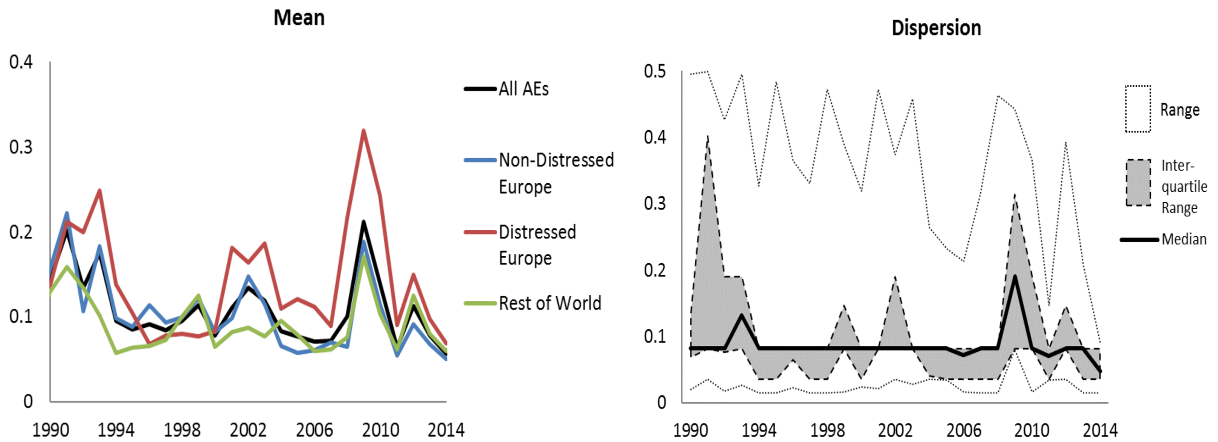
Figure 6. Crisis Probabilities at the Individual Country Level



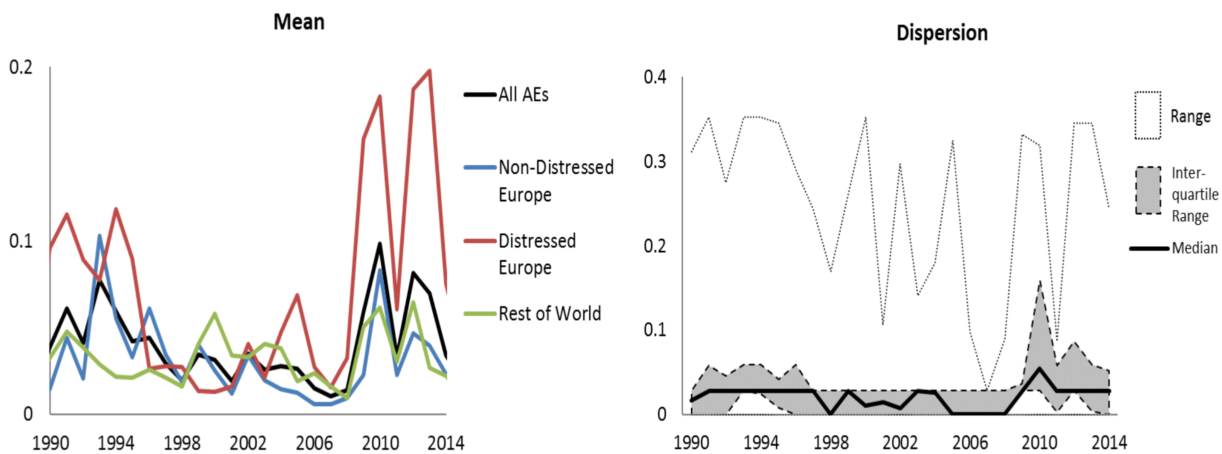
Source: Authors' calculations.
Note: HH = Household balance sheet; MT = Medium term.

Figure 7. AE Crisis Probabilities at the Global and Regional Levels

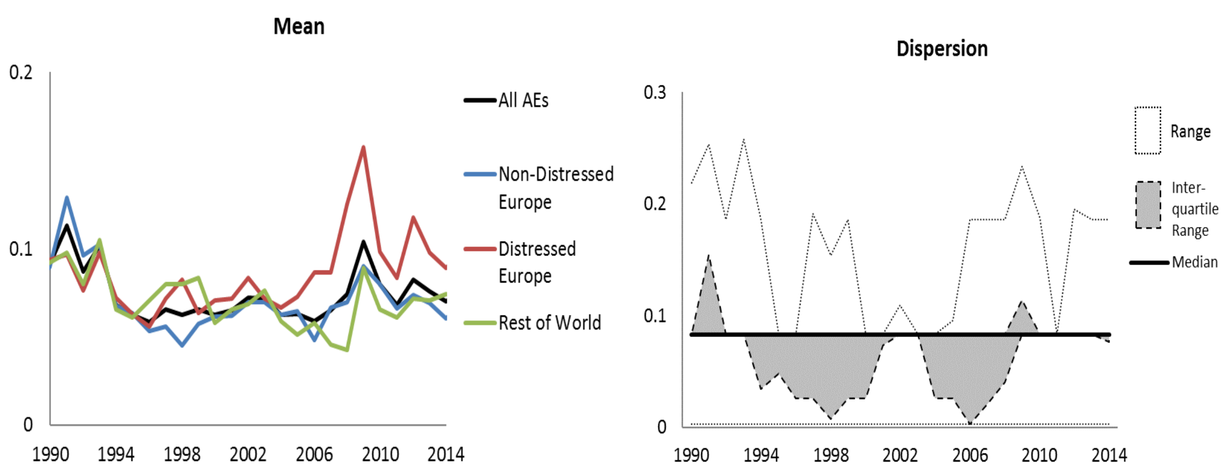
a) Growth Probabilities



b) Fiscal Probabilities

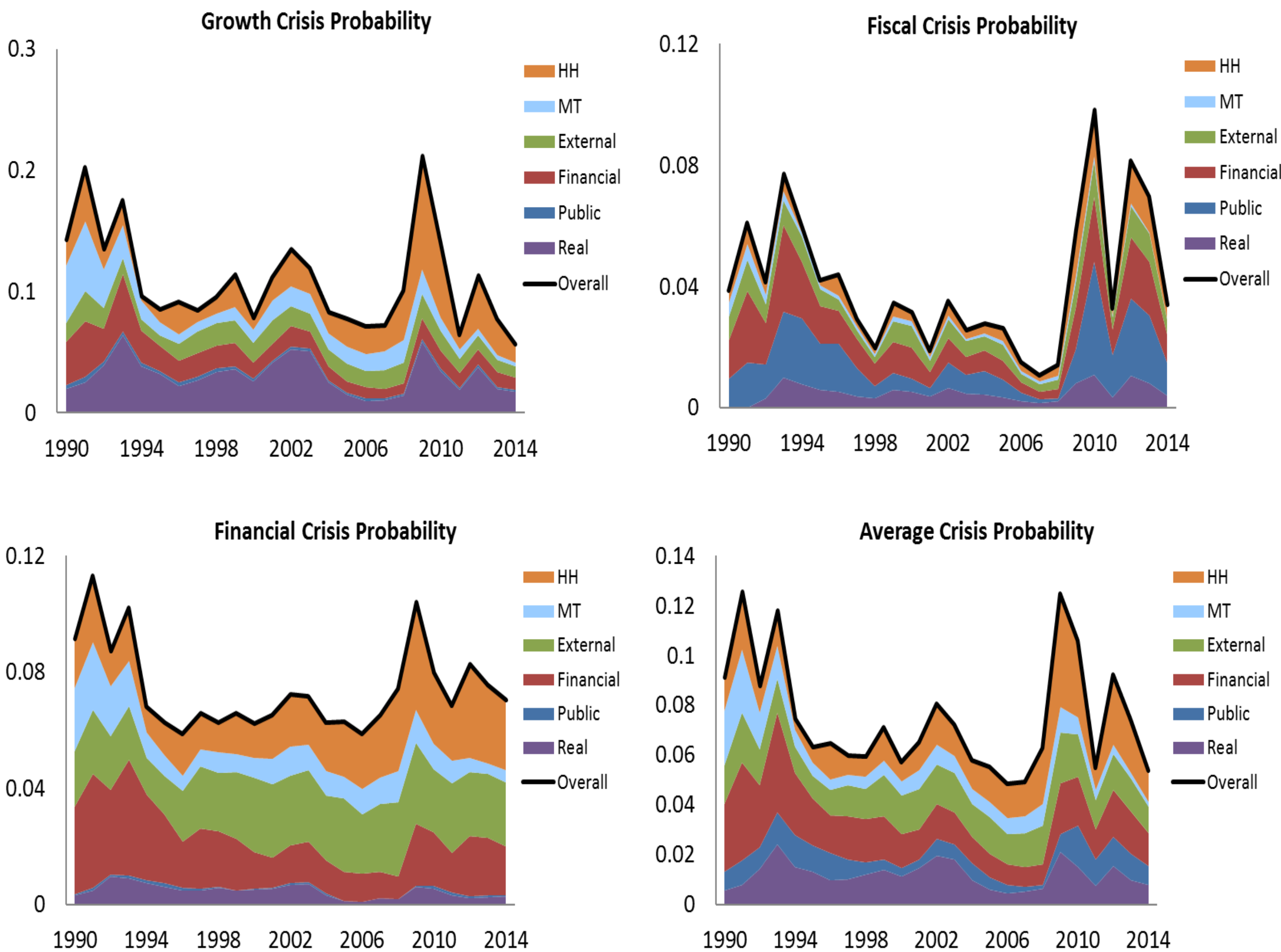


c) Financial Probabilities



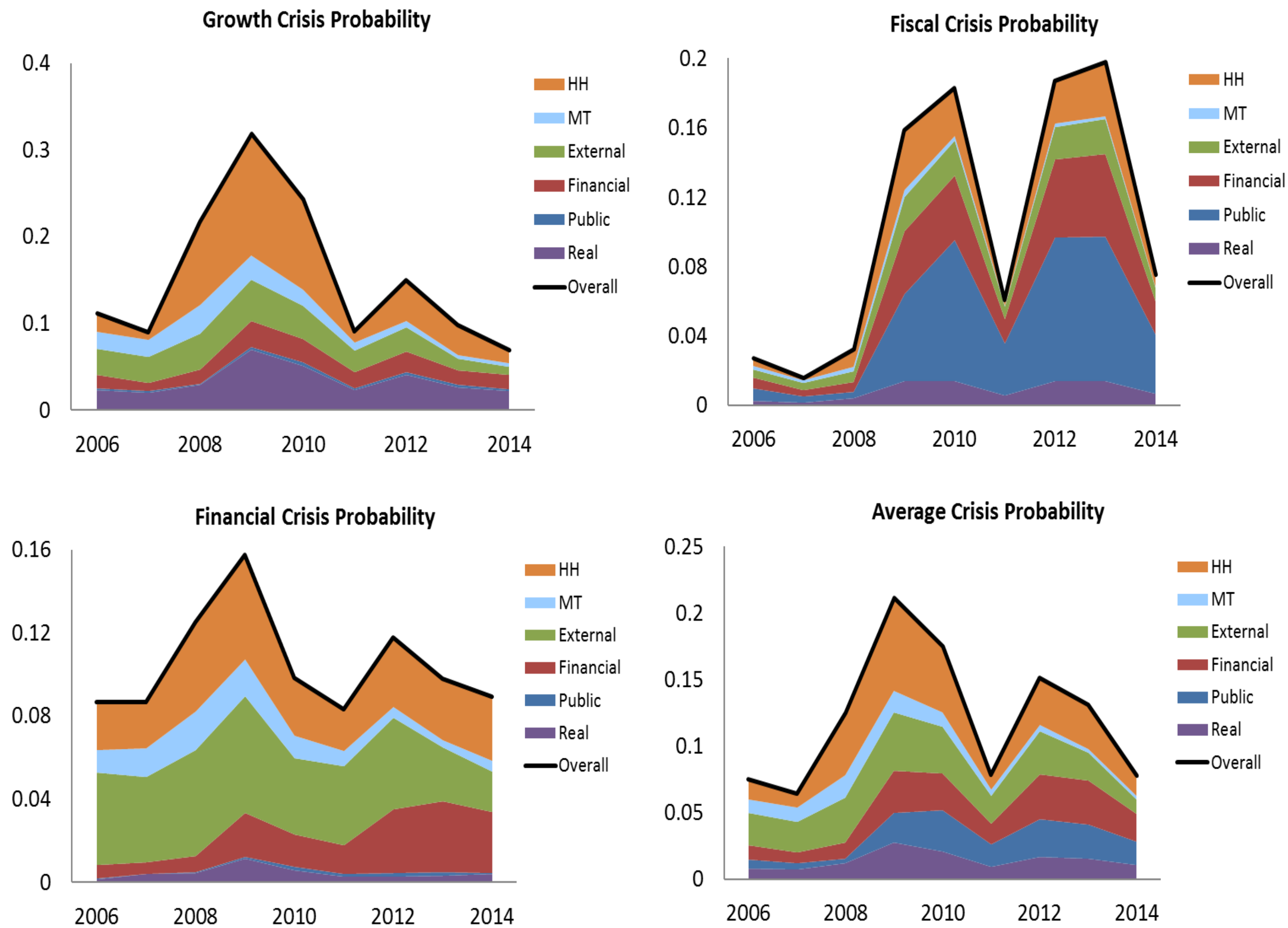
Source: Authors' calculations.
Note: Appendix Table A1 contains the list of countries in the AE sample and their breakdown by region.

Figure 8. Sectoral Contributions to Global Crisis Probabilities in AEs



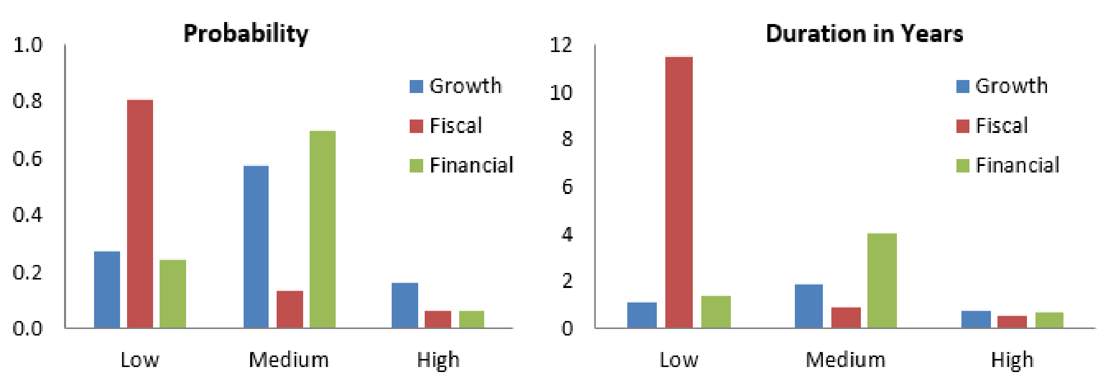
Source: Authors' calculations.
Note: HH = Household balance sheet; MT = Medium term.

Figure 9. AEs in Distressed Europe during the Global Crisis



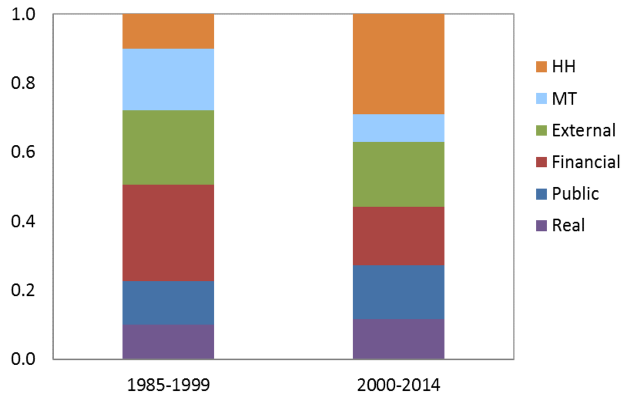
Source: Authors' calculations.
Note: HH = Household balance sheet; MT = Medium term.

Figure 10. Probability and Duration of Ratings for AEs, 1985–2017



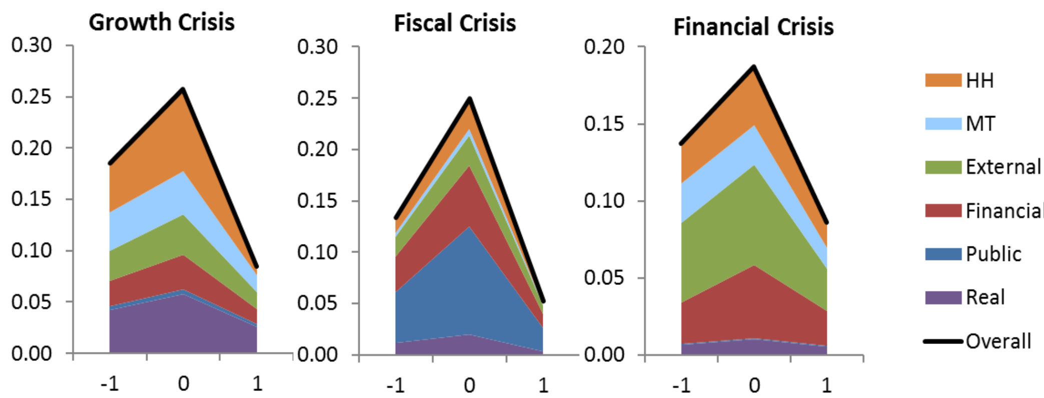
Source: Authors' calculations.
Note: Probabilities and durations are calculated using the transition matrices in table 4.

Figure 11. Sources of High Ratings for AEs



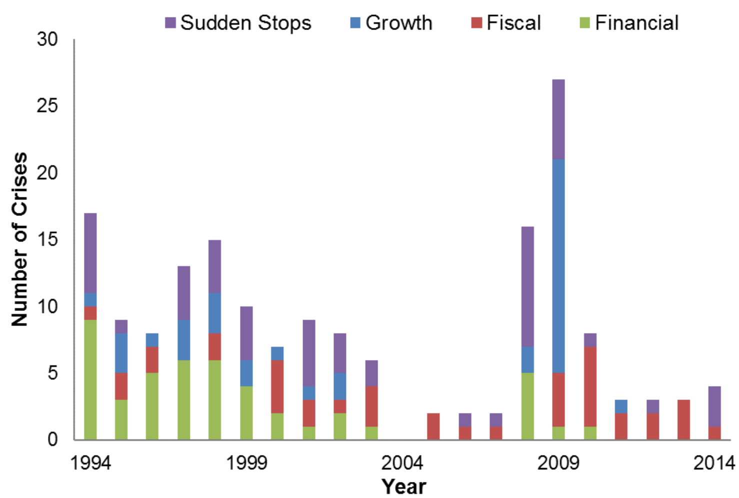
Source: Authors' calculations.
Note: HH = Household balance sheet; MT = Medium term.

Figure 12. Transition from High to Medium/Low Rating for AEs, 1985–2017



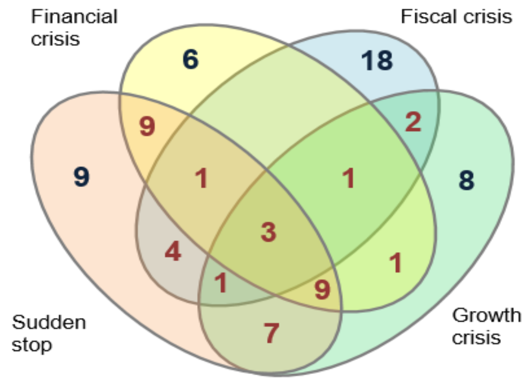
Source: Authors' calculations.
Note: HH = Household balance sheet; MT = Medium term. Date 0 is the last year with a high rating before a downgrade in rating; Dates -1 and 1 are years before and after.

Figure 13. Crisis Incidence in EMs



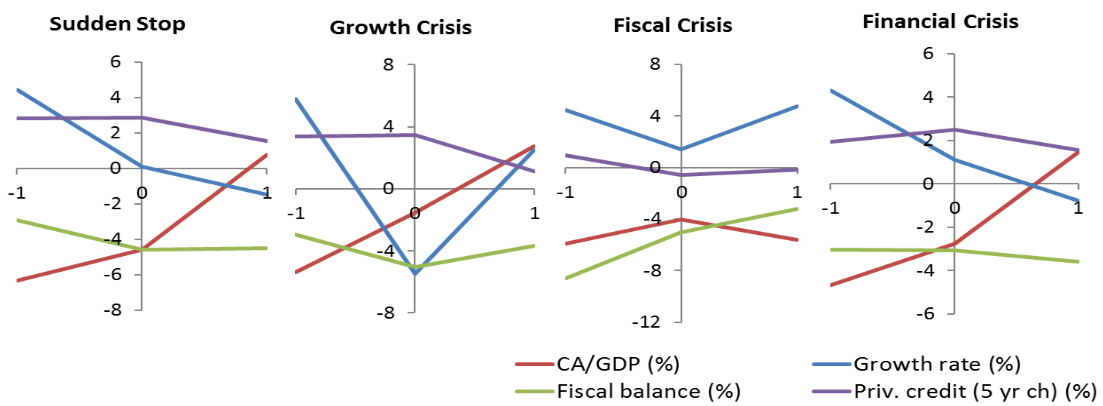
Source: Laeven and Valencia (2012) and authors' calculations.
Note: Data on financial crisis incidence is available only up to 2010.

Figure 14. Overlap of EM Crises, 1994–2010



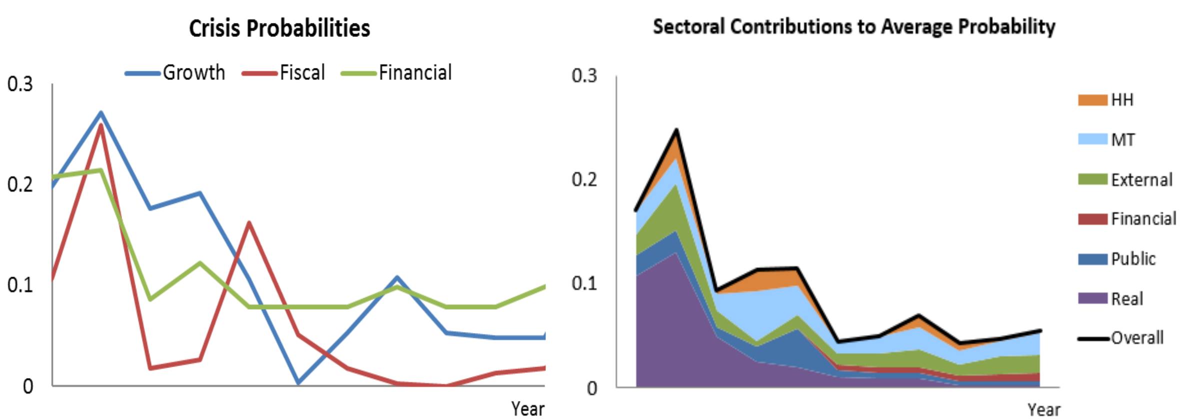
Source: Authors' calculations.
Note: Crises in the same or adjacent years are regarded as coincident.

Figure 15. Macroeconomic Dynamics Around EM Crises, 1994–2010



Source: Authors' calculations.
Note: Date 0 is crisis year; Dates -1 and 1 are years before and after.

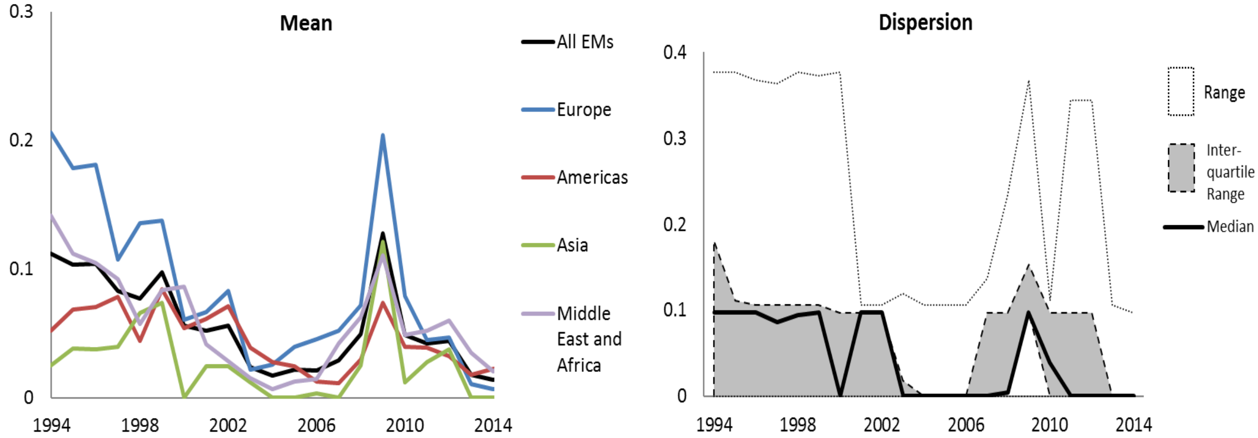
Figure 16. Crisis Probabilities at the Individual Country Level



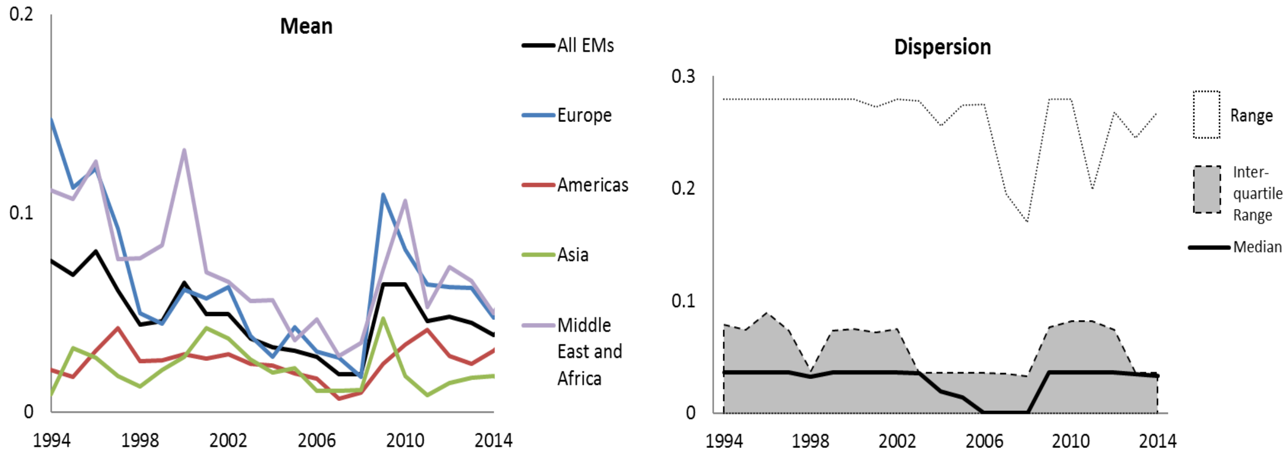
Source: Authors' calculations.
Note: HH = Household balance sheet; MT = Medium term.

Figure 17. EM Crisis Probabilities at the Global and Regional Levels

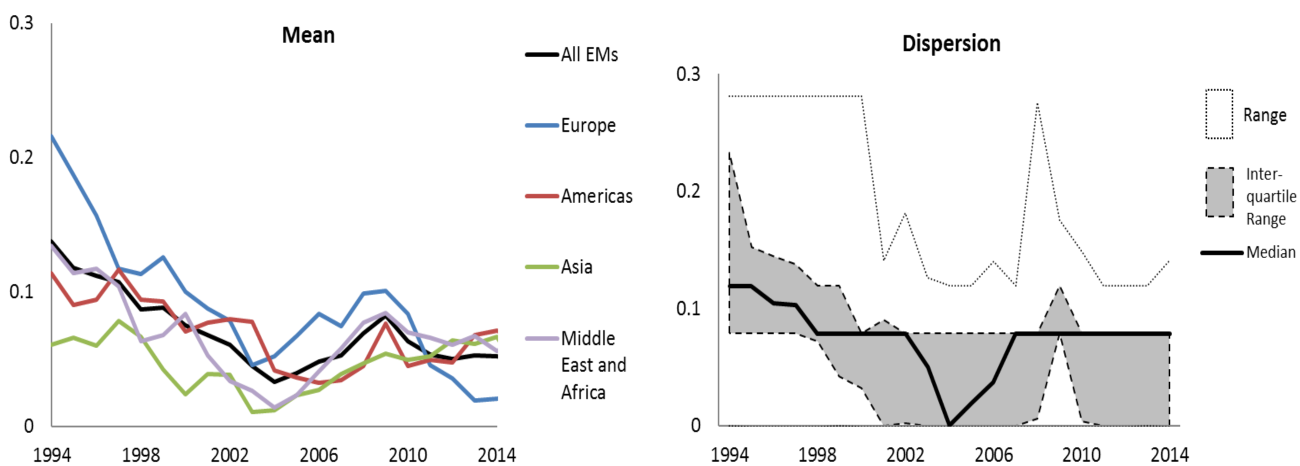
a) Growth Probabilities



b) Fiscal Probabilities

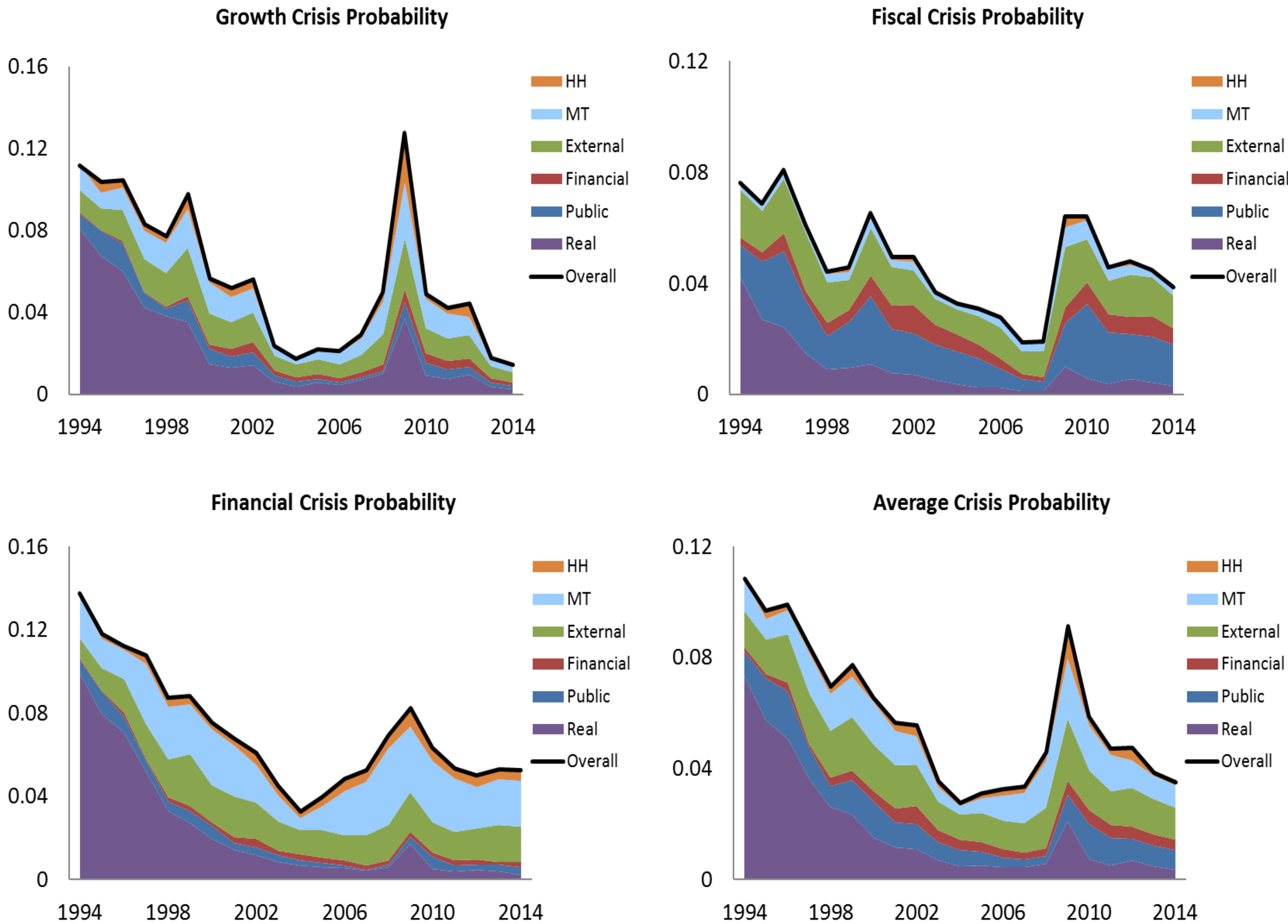


c) Financial Probabilities



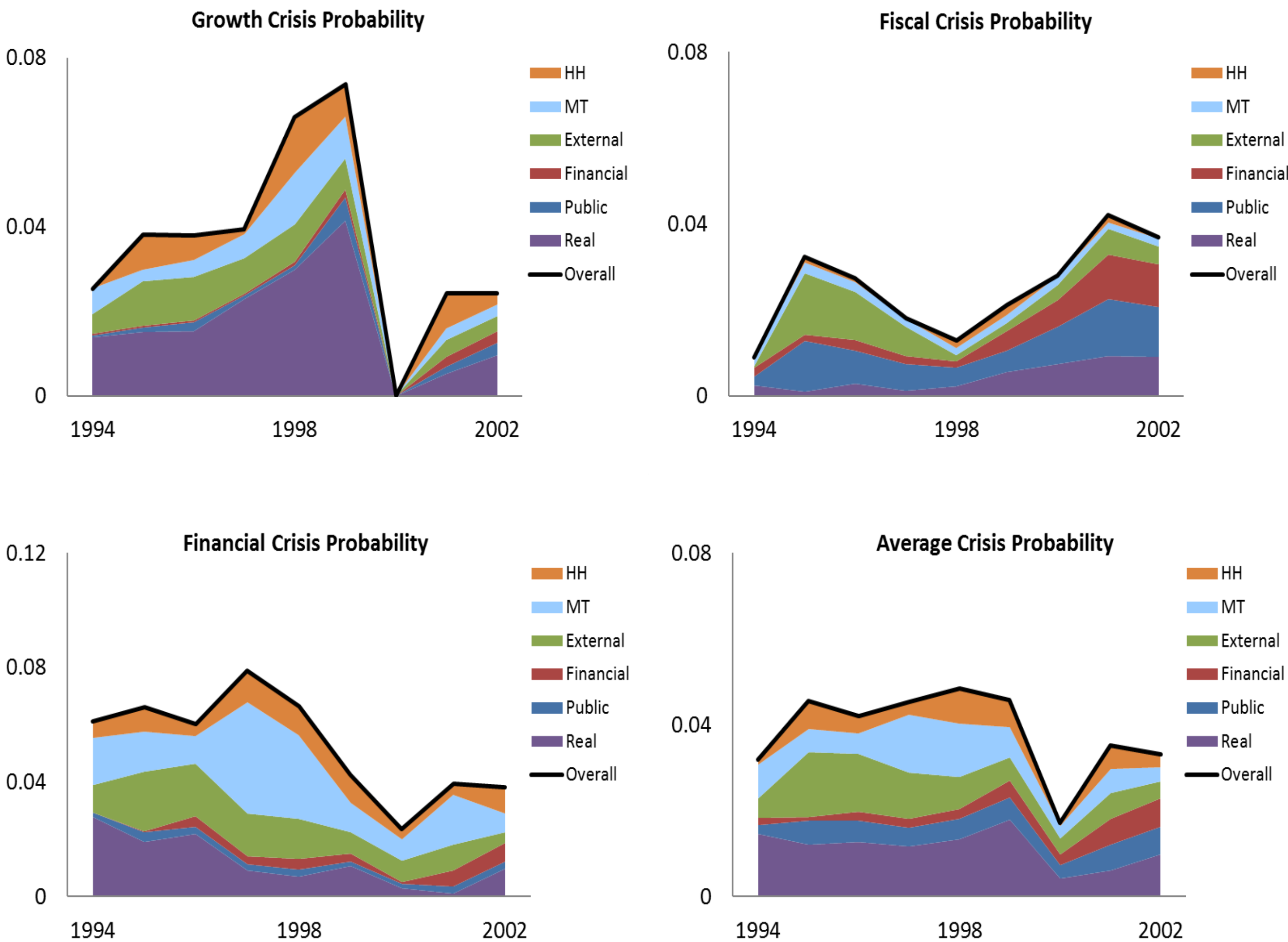
Source: Authors' calculations.
Note: Appendix Table A2 contains the list of countries in the EM sample and their breakdown by region.

Figure 18. Sectoral Contributions to Global Crisis Probabilities in EMs



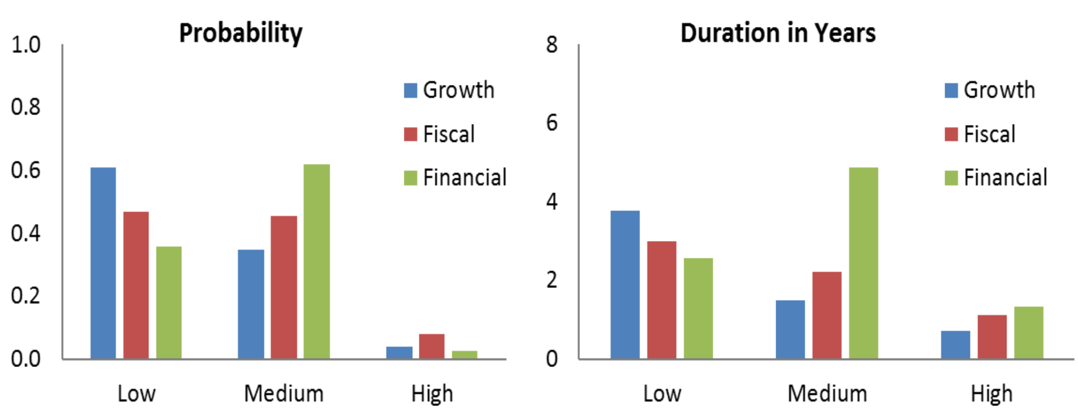
Source: Authors' calculations.
Note: HH = Household balance sheet; MT = Medium term.

Figure 19. EMs in Asia during the Asian Crisis



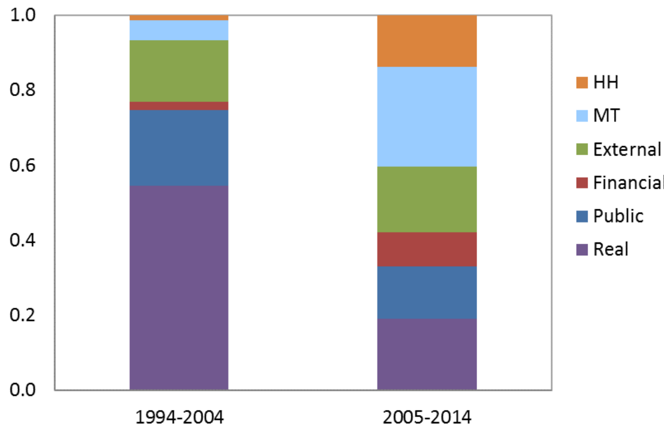
Source: Authors' calculations.
Note: HH = Household balance sheet; MT = Medium term.

Figure 20. Probability and Duration of Ratings for EMs, 1994–2017



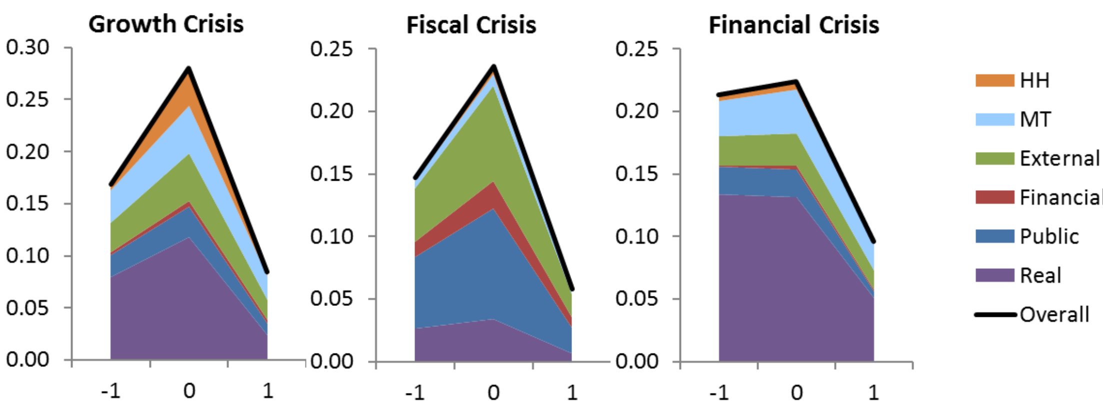
Source: Authors' calculations.
Note: Probabilities and durations are calculated using the transition matrices in table 9.

Figure 21. Sources of High Ratings for EMs



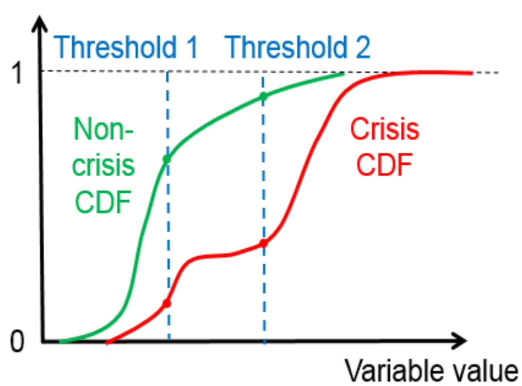
Source: Authors' calculations.
Note: HH = Household balance sheet; MT = Medium term.

Figure 22. Transition from High to Medium/Low Rating for EMs, 1994–2017



Source: Authors' calculations.
Note: HH = Household balance sheet; MT = Medium term. Date 0 is the last year with a high rating before a downgrade in rating; Dates -1 and 1 are years before and after.

Figure 23. Threshold Robustness



Source: Authors' illustration.

Table 1. Crisis Frequency in AEs

	Full sample 1985-2010	Estimation sample 1985-2007	Post-crisis sample 2008-2010
Growth crisis	0.08	0.05	0.30
Fiscal crisis	0.03	0.02	0.06
Financial crisis	0.04	0.02	0.16

Note: Frequencies calculated over non-missing crisis observations.

Table 2. Data used in the AE Crisis Probability Models

Variable	Source 1/	Variable	Source 1/
Crisis calculations		**Public sector**	
Real GDP	WEO	General government balance/GDP	WEO
Cyclically-adjusted primary deficit/GDP	OECD EO	Primary gap/GDP 3/	WEO
		Public debt/GDP 3/	WEO
Financial crisis dummy	LV	**Financial sector**	
		Capital adequacy ratio	Fitch
Medium-term		Return on bank assets	Fitch
Private sector credit growth (5-yr)	WB WDI	Non-performing loans	Fitch
Housing price growth (5-yr) 2/	OECD HPD, GPG	**Household balance sheets**	
Stock price growth (5-yr) 2/	Bloomberg	Housing price acceleration 2/	OECD HPD, GPG
Construction sector contribution to GDP growth (5-yr)	OECD	Stock price acceleration 2/	Bloomberg
		Household liabilities/GDP	OECD EO
Financial sector contribution to GDP growth (5-yr)	OECD	Interaction: Household liab./GDP *Housing price growth (5-yr)	OECD EO, OECD HPD
		Interaction: Household liab./GDP *Housing price acceleration	OECD EO, OECD HPD
Near-term			
Real sector		**External Sector**	
Black-Scholes-Merton default prob.	CVU	Current account/GDP	WEO
Return on corporate assets	CVU	External debt/GDP	EWN, WEO
Price to earnings ratio	CVU	External debt/exports	EWN, WEO
Interest coverage ratio	CVU	Deviation from EBA norm	EBA
Inflation	WEO		

1/ WEO = IMF World Economic Outlook, CVU = IMF Corporate Vulnerability Utility, EBA = IMF External Balance Assessment, WB WDI = World Bank World Development Indicators, OECD = OECD Website, OECD EO = OECD Economic Outlook, OECD HPD = OECD Housing Prices Database, GPG = Global Property Guide, LV = Laeven and Valencia (2012), EWN = External Wealth of Nations dataset, updated version of Lane and Milesi-Ferretti (2007).
2/ Nominal housing and stock prices are used in the VEA because they interact with nominal contracts held on financial sector and household balance sheets. Price acceleration is defined as the change in the annual growth rate of the price.
3/ Primary gap/GDP and public debt/GDP variables are used only for the fiscal crisis modules of the VEA.

Table 3. Sum of Errors for AEs

	Full sample 1985-2010	Estimation sample 1985-2007	Post-crisis sample 2008-2010
Growth crisis	0.46	0.39	0.70
Fiscal crisis	0.26	0.28	0.28
Financial crisis	0.57	0.24	0.96

Note: Errors calculated over non-missing crisis observations.

Table 4. Transition Matrices for AEs, 1985–2017

Growth

Rating at t		Rating at t+1: L	M	H
	L	0.52	0.46	0.02
	M	0.21	0.65	0.15
	H	0.06	0.51	0.43

Fiscal

Rating at t		Rating at t+1: L	M	H
	L	0.92	0.06	0.03
	M	0.41	0.46	0.13
	H	0.22	0.43	0.35

Financial

Rating at t		Rating at t+1: L	M	H
	L	0.58	0.42	0.00
	M	0.14	0.80	0.05
	H	0.01	0.58	0.41

Table 5. Crisis Frequency in EMs

	Estimation sample 1994-2010	Pre-crisis sample 1994-2007	Post-crisis sample 2008-2010
Sudden stop	0.05	0.04	0.10
Growth crisis	0.04	0.03	0.11
Fiscal crisis	0.04	0.03	0.06
Financial crisis	0.05	0.05	0.05

Note: Frequencies calculated over non-missing crisis observations.

Table 6. Data used in the EM Crisis Probability Models

Variable	Source 1/	Variable	Source 1/
Crisis calculations		**Public sector**	
Real GDP	WEO	General government balance/GDP	WEO, CSD
Cyclically-adjusted primary deficit/GDP	FAD FP	Primary gap/GDP 3/	WEO, CSD
		Public debt/GDP 3/	WEO, CSD
Financial crisis dummy	LV	EMBI sovereign spread	Bloomberg
		Financial sector	
Medium-term		Capital adequacy ratio	CSD
Private sector credit growth (5-yr)	WB WDI	Return on bank assets	CSD
Housing price growth (5-yr) 2/	OECD HPD, GPG	Non-performing loans	CSD
Stock price growth (5-yr) 2/	Bloomberg	**Household balance sheets**	
REER growth (5-yr)	INS	Housing price acceleration 2/	OECD HPD, GPG
		Stock price acceleration 2/	Bloomberg
Near-term		**External Sector**	
Real sector		Current account/GDP	WEO
Black-Scholes-Merton default prob.	CVU, CSD	External debt/GDP	WEO, CSD
Return on corporate assets	CVU, CSD	External debt/exports	WEO, CSD
Price to earnings ratio	CVU, CSD	Deviation from EBA norm	EBA
Interest coverage ratio	CVU, CSD	REER acceleration	INS
Inflation	WEO	Absolute oil balance/GDP	WEO

1/ WEO = IMF World Economic Outlook, CSD = IMF Common Surveillance Database, CVU = IMF Corporate Vulnerability Utility, EBA = IMF External Balance Assessment, FAD FP = FAD Fiscal Policy and Surveillance Database, INS = IMF Information Notice System, WB WDI = World Bank World Development Indicators, OECD HPD = OECD Housing Prices Database, GPG = Global Property Guide, LV = Laeven and Valencia (2012).
2/ Real housing and stock prices are used in the VEE because nominal data suffer from structural breaks in inflation. Price acceleration is defined as the change in the annual growth rate of the price.
3/ Primary gap/GDP and public debt/GDP variables are used only for the fiscal crisis modules of the VEE.

Table 7. Sum of Errors for EMs

	Estimation sample 1994-2010	Pre-crisis sample 1994-2007	Post-crisis sample 2008-2010
Growth crisis	0.42	0.39	0.53
Fiscal crisis	0.37	0.45	0.22
Financial crisis	0.52	0.55	0.39

Note: Errors calculated over non-missing crisis observations.

Table 8. Sum of Errors for Truncated EM Model

	Full sample 1994-2010	Estimation sample 1994-2007	Post-crisis sample 2008-2010
Growth crisis	0.63	0.40	0.76
Fiscal crisis	0.39	0.34	0.51
Financial crisis	0.59	0.58	0.67

Note: Errors calculated over non-missing crisis observations.

Table 9. Transition Matrices for EMs, 1994–2017

Growth

		Rating at t+1 L	M	H
Rating at t	L	0.79	0.20	0.01
	M	0.35	0.60	0.05
	H	0.08	0.49	0.42

Fiscal

		Rating at t+1 L	M	H
Rating at t	L	0.75	0.23	0.02
	M	0.25	0.69	0.06
	H	0.08	0.39	0.53

Financial

		Rating at t+1 L	M	H
Rating at t	L	0.72	0.27	0.00
	M	0.16	0.83	0.02
	H	0.07	0.36	0.57

Table A1. AE Sample of Countries and Assigned Regions

Country	Region	Country	Region
Australia	Rest of World	Japan	Rest of World
Austria	Non-Distressed Europe	Korea, Republic of	Rest of World
Belgium	Non-Distressed Europe	Luxembourg	Non-Distressed Europe
Canada	Rest of World	Malta	Non-Distressed Europe
China,P.R.:Hong Kong	Rest of World	Netherlands	Non-Distressed Europe
Cyprus	Distressed Europe	New Zealand	Rest of World
Czech Republic	Non-Distressed Europe	Norway	Non-Distressed Europe
Denmark	Non-Distressed Europe	Portugal	Distressed Europe
Estonia	Non-Distressed Europe	Singapore	Rest of World
Finland	Non-Distressed Europe	Slovak Republic	Non-Distressed Europe
France	Non-Distressed Europe	Slovenia	Non-Distressed Europe
Germany	Non-Distressed Europe	Spain	Distressed Europe
Greece	Distressed Europe	Sweden	Non-Distressed Europe
Iceland	Distressed Europe	Switzerland	Non-Distressed Europe
Ireland	Distressed Europe	United Kingdom	Non-Distressed Europe
Israel	Rest of World	United States	Rest of World
Italy	Distressed Europe		

Note: Distressed Europe captures those European countries which received external financial assistance from the IMF, EFSF/ESM, and/or ECB SMP. We classify SMP participation following ECB (2013).

Table A2. EM Sample of Countries and Assigned Regions

Country	Region	Country	Region
Albania	Europe	Kazakhstan	Middle East and Africa
Algeria	Middle East and Africa	Latvia	Europe
Angola	Middle East and Africa	Lebanon	Middle East and Africa
Argentina	Americas	Lithuania	Europe
Armenia	Middle East and Africa	Macedonia	Europe
Azerbaijan	Middle East and Africa	Malaysia	Asia
Bahamas	Americas	Mauritius	Middle East and Africa
Belarus	Europe	Mexico	Americas
Bosnia&Herzegovina	Europe	Morocco	Middle East and Africa
Brazil	Americas	Pakistan	Middle East and Africa
Bulgaria	Europe	Panama	Americas
Chile	Americas	Peru	Americas
China	Asia	Philippines	Asia
Colombia	Americas	Poland	Europe
Costa Rica	Americas	Romania	Europe
Croatia	Europe	Russia	Europe
Dominican Republic	Americas	Serbia	Europe
Ecuador	Americas	South Africa	Middle East and Africa
Egypt	Middle East and Africa	Sri Lanka	Asia
El Salvador	Americas	Thailand	Asia
Georgia	Middle East and Africa	Tunisia	Middle East and Africa
Guatemala	Americas	Turkey	Europe
Hungary	Europe	Ukraine	Europe
India	Asia	Uruguay	Americas
Indonesia	Asia	Venezuela	Americas
Jamaica	Americas	Vietnam	Asia
Jordan	Middle East and Africa		