Chapter 4
On the Evaluation of Early Warning Models for Financial Crises

Lucia Alessi
European Central Bank, Germany

Carsten Detken
European Central Bank, Germany

Silviu Oprică
Goethe-University Frankfurt, Germany

ABSTRACT

Early Warning Models (EWMs) are back on the policy agenda. In particular, accurate models are increasingly needed for financial stability and macro-prudential policy purposes. However, owing to the alleged poor out-of-sample performance of the first generation of EWMs developed in the 90's, the economic profession remains largely unconvinced about the ability of EWMs to play any important role in the prediction of future financial crises. The authors argue that a lot of progress has been made recently in the literature and that one key factor behind the prevailing skepticism relates to the basic evaluation metrics (e.g. the noise-to-signal ratio) traditionally used to evaluate the predictive performance of EWMs, and in turn to select benchmark models. This chapter provides an overview of methodologies (e.g. the (partial) Area Under the Receiver Operating Characteristic curve and the (standardized) Usefulness measure) better suitable for measuring the goodness of EWMs and constructing optimal monitoring tools.

INTRODUCTION

Owing to the global financial crisis, Early Warning Models (EWMs) are back on the policy agenda. In particular, we argue that EWMs for financial crises are increasingly needed to support policy decisions in the macro-prudential field and several key institutions are indeed designing new tools for the early detection of vulnerabilities. This has triggered a renewed interest in academic research on the topic. On the other hand, owing to the alleged poor out-of-sample performance of the

first generation of EWMs developed in the 90's, large parts of the economic profession remain unconvinced about the ability of EWMs to predict financial crises and therefore the usefulness of EWMs to support the policy process.

There are several good reasons why the academic profession has been skeptical about the contribution of EWMs to predict financial crises. Eichengreen (2003) states that EWMs have difficulties in dealing with structural relationships which interact in non-linear and state-contingent ways. E.g. an asset price misalignment should be interpreted differently if the overvalued assets are held by leveraged and interconnected agents and the vulnerability is unlikely to be increasing proportionally with the degree of overvaluation. He further notes that complex systems, like the financial system, often have multiple equilibria, which are sensitive to small perturbations. And finally he reminds us of the circularity of forecasts/warnings, which is also known as Goodhart's Law.[1] The latter states that any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes. If e.g. macro-prudential authorities would use EWMs and act upon the signals in an appropriate way, crises would be prevented and no future econometrician would be able to detect the early warning properties of any indicator or model in a reduced form type of analysis. In summary, Eichengreen highlights that the role of fundamentals in predicting financial crises might be difficult to detect by means of (standard) empirical analysis. Furthermore, the early warning literature suffered in the past from frequent out-of-sample failures, partly based on in-sample over-fitting and variable selection bias. What this means is that often EWMs have been fit to explain a particular crisis or set of similar crises, e.g. the Asian current account crisis in 1997, but the indicators and models were selected after the fact. The models and performance measures were not derived in real time and thus oversold the capabilities of such models to predict future crises ex-ante.[2] Policy makers and academic jour-

nal referees alike have not been amused by what often appeared like exercises in pure data-mining.[3]

More recently though, a lot of progress has been made on many fronts. Out-of-sample validation methods are increasingly used to select the best performing models. This means that the estimation and evaluation samples are split in order to obtain realistic performance statistics.[4] Objective methods are being used to avoid in-sample variable selection bias and obtain concise model specifications, which perform well out-of-sample.[5] Real vintage data are not available in most cases, but pseudo real time exercises are being performed, e.g. by lagging data in a way that at each point in time only data are used, which would have been truly available at the time of the forecast. If trends or deviations from trends (gaps) are being used, they are calculated recursively to comply with the same real time concept.[6] EWMs are also now geared towards the use of policy instruments by aiming to predict only those types of crises, which a specific policy instrument is meant to prevent or alleviate. This makes the signals received more useful from the policy maker's perspective.[7] A key element is also that EWMs usually do not predict the outbreak of a crisis, which is nearly impossible to do, as very small random events can trigger a crisis. The trigger might even be unrelated to fundamentals (see Eichengreen, 2003), while the underlying vulnerability is not. EWMs nowadays identify a vulnerable state, which would tend to lead to a crisis within a certain time span, e.g. within the coming 3 years. This gives more relevance to the prediction and would allow policy makers to take relevant action ahead of the crisis.

And finally, we argue that one key factor behind the dismal performance of EWMs is related to the basic evaluation metrics (e.g. the noise-to-signal ratio) traditionally used to evaluate their predictive performance. This in turn led to a suboptimal selection of benchmark models. In this respect, this chapter provides an overview of available methodologies suitable for measuring the goodness of EWMs. Some of them are

Related Content

The Impact of Ownership Structure on Firm Performance: Evidence from Pakistan
Qaiser Rafique Yasserand Abdullah Al Mamun (2016). *International Journal of Corporate Finance and Accounting (pp. 36-54).*
www.irma-international.org/article/the-impact-of-ownership-structure-on-firm-performance/164986

Linear Integer Programming Techniques for Portfolio Design Problem: Linear Programming Approaches
(2020). *Algorithms for Solving Financial Portfolio Design Problems: Emerging Research and Opportunities (pp. 80-102).*
www.irma-international.org/chapter/linear-integer-programming-techniques-for-portfolio-design-problem/250521

Global Trends in Domestic savings and Longevity- Implications for India
Thilak Venkatesanand Venkataraman R (2019). *International Journal of Corporate Finance and Accounting (pp. 18-27).*
www.irma-international.org/article/global-trends-in-domestic-savings-and-longevity--implications-for-india/234335

Empirical Analysis of Dependence on Capital Structure with Respect to Its Impact on the Financial Performance of Domestic Non Financial Listed Stock Exchange Companies Operating in the Public Sector of Pakistan
Syed Jawad Hussain Shahzadand Memoona Kanwal (2016). *International Journal of Corporate Finance and Accounting (pp. 58-76).*
www.irma-international.org/article/empirical-analysis-of-dependence-on-capital-structure-with-respect-to-its-impact-on-the-financial-performance-of-domestic-non-financial-listed-stock-exchange-companies-operating-in-the-public-sector-of-pakistan/174421

The Adoption Process of Payment Cards: An Agent-Based Approach
Biliana Alexandrova-Kabadjova, Sara G. Castellanos Pascacioand Alma L. García-Almanza (2013). *Simulation in Computational Finance and Economics: Tools and Emerging Applications (pp. 1-28).*
www.irma-international.org/chapter/adoption-process-payment-cards/68688