Could we have seen it coming? Towards an early warning system for asylum applications in the EU*

Emily Barker [†] Jakub Bijak[‡]

September 21, 2023

Abstract

Forecasting large changes in the numbers of asylum applications, an element of so-called asylum 'crises', is challenging. In this paper, we present a model that shows that the warning signs of a crisis could appear in publicly-available macroeconomic, geopolitical, and demographic data sources. Our aim is to propose and test an early warning system for asylum applications in the EU+ that would be easy to use, effective and interoperable for policy makers, and that would give sufficient advance warning that authorities can be prepared for an increase in the number of asylum applications up to six months in advance for two of the most prominent asylum flows from the recent decade, involving people fleeing the wars in Syria and in Ukraine.

JEL Classification: C52, C53, E27, F22, F42, J11
Keywords: Demographic forecasting, Evaluating forecasts, Model selection, Probability forecasting, Scenarios, Uncertainty

*We are very grateful to Marta Bivand Erdal, Rainer Münz and Teddy Wilkin for their comments that helped improve an earlier draft; to Peter WF Smith and Jason Hilton for methodological suggestions, and to seminar participants at the RES and IMISCOE 2023 conferences. This research is part of the QuantMig research project that is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 870299. All the remaining errors and inaccuracies are ours. This document reflects the author's view and the Research Executive Agency of the European Commission are not responsible for any use that may be made of the information it contains.

†The University of Southampton, Email: E.R.Barker@soton.ac.uk
‡The University of Southampton, Email: J.Bijak@soton.ac.uk

1 Introduction

There is large uncertainty around international migration flows in general; particularly for the numbers of asylum seekers related to high-impact events. These 'crises', as they are often perceived, involve large numbers of asylum seekers fleeing their home countries due to politically-motivated factors, such as war, state violence, or persecution. The magnitude and impact of these flows vary in size and intensity; they can be related to single countries of origin, such as Afghanistan, Bosnia and Herzegovina, Kosovo or Ukraine, or whole regions, e.g. Middle Eastern and North African countries since the Arab Spring. In the latter group, while individual countries were affected by different political events, it was the war in Syria that has become almost synonymous with asylum.

Recently, there has been a considerable discussion whether the 'crisis' as experienced by host countries is related to migration *per se* or rather to its governance (Crawley, 2016; Carastathis et al., 2018; Bijak and Czaika, 2020; Sahin-Mencutek et al., 2022). The surge in asylum applications mainly from Syria in 2015–16 resulted in many practical challenges in the destination countries, which could have arisen from both the number of people seeking protection, and from how the governments dealt with the sudden increases in the inflows. In this paper, we use the term 'crisis' to refer to an unforeseen arrival of large numbers of asylum seekers in the European Union (EU) and its neighbourhood.[1] The term 'crisis' is used in a generic sense, as possibly related both to the asylum processes and their governance aspects, while realising that the governance response can be pivotal in either addressing or exacerbating challenges related to sudden increase in the numbers of migrants, including asylum seekers.

The early warnings systems for asylum applications are intended to alert policy makers – in our case, in the EU – about upcoming increases in the numbers of asylum seekers, which are not typically foreseeable with traditional methods of analysis. The insights from an early warning system (hereafter, EWS) can allow policy makers to improve the management of asylum-related migration at different decision levels by providing a more timely and potentially better targeted operational response at subsequent stages of the process. The main aim of warning decision makers early is to give them time for support

[1]For practical reasons, related to the availability of data, we use the number of asylum applications as an imperfect proxy measure – and also as a variable of interest in its own right.

networks to be put in place to cope with an increase in the number of asylum applications, and if required to be prepared for a policy change.

In practice, in the light of high uncertainty, before even considering the questions of migration governance, it is understandable that migration 'crises' seem ill-prepared for. Researchers and policy makers alike have, however, learned some lessons from the 2015–16 asylum crisis driven by inflows from Afghanistan, Iraq and notably Syria, in particular with research shifting to the early signal detection (e.g. Napierała et al., 2022; Carammia et al., 2022), and with policy responses moving towards preparedness. The latter involves the EU Blueprint on preparedness and crisis management mechanism in the area of migration[2]. Migration policies can be slow to change, but as has been seen during the 2015–16 crisis, and even more so in the (rapid) policy shift following the 2022 invasion of Ukraine, swift action can be put in place when needed.

The literature on early warning systems in the context of migration and asylum remains relatively limited. The policy and research interest in this area was piqued following the rise of Syrian migration to Europe of 2015–16. Yet, there has not been a large development or optimisation of models which could be tested; instead, the formal work in this area remains largely in the prototype phase. The research presented in this paper builds on the existing work by Napierała et al. (2022) and Carammia et al. (2022) who provided examples of early warning models for migration, with the former looking at detection of change-points in asylum series, and the latter focused on forecasting of migration flows with 'big data'. Initial research on the 2022 Ukraine crisis by Juric (2022), used Google Trends search data focusing on migration planning terms, noting an important limitation related to searches made in the Russian language as well as in the Ukrainian language.

We start by defining 'crisis' through binary indicators, determined by the volumes and growth rates in the numbers of asylum applications. We employ selected 'big data' sources as leading predictors of the crisis at different horizons, of up to six months, for two recent case studies of asylum applications from Syria and Ukraine. As the models uses a large selection of variables, Least Absolute Shrinkage and Selection Operator (LASSO) is applied to select the variables to be included in the EWS models. For both case studies,

[2]Commission Recommendation (EU) 2020/1366 of 23 September 2020 on an EU mechanism for preparedness and management of crises related to migration, OJ L 317, 1.10.2020, p. 26–38, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2020:317:TOC.

we assess a number of models at different horizons. The models presented are designed to trigger alerts up to six months prior to the surge in asylum applications, which should enhance the decision making and preparedness. We use contemporaneous models as a comparison for the forecasts. The applications of our models are illustrated for two case studies: the Syrian 'asylum crisis' of 2015–16 and the recent (and ongoing) migration implications of the war and humanitarian crisis in Ukraine.

The remainder of this paper is organised as follows: Section 2 introduces the methodology behind our early warning models, as well as operationalising the definition of a 'crisis'; Section 3 presents the applications of the model for the Syrian and Ukraine case studies; with the results and discussion in Section 4; and Section 5 discusses the practical and policy implication of the results, and offers a conclusion.

2 Methodology

In this section, we explain the model employed for the early warning analysis and summarise the data sources used for the various input variables. We also briefly discuss problem with defining a 'crisis event' for use in early warning models.

2.1 Model description

In an early warning model with a binary response variable, for each period in the observation window, the binary variable takes a value of 0 to indicate no crisis or 1 for a crisis. The model thus estimates a probability, $\widehat{Pr}$, that a crisis will occur, which will trigger an early warning if this probability is greater than some threshold value, c. There are two types of error, where a crisis is misidentified: the occurrence of a false negative (Type I error), occurs when the probability does not meet the threshold level (no signal) but a crisis occurs. A false positive (Type II error) occurs when the probability of a crisis is greater than the threshold level (signal), but a crisis does not occur. Kauppi and Saikkonen (2008) used a dynamic binary probit approach as a general framework for defining four dichotomous models to evaluate historical data and to forecast potential crises. In its general form, such dichotomous model can be written as:

$$P_{t-1}(y_t = 1) = F(\pi_t) = F(\beta' \mathbf{X_{t-1}} + \Sigma_{j=1}^{p} \gamma_j y_{t-j} + \Sigma_{j=1}^{q} \eta_j \pi_{t-j}), \tag{1}$$

where y_t is the dichotomous variable at time t, π_t is the corresponding latent variable, $\mathbf{X_{t-1}}$ is a kx1 vector of k explanatory variables, and β is a kx1 vector that responds to the coefficients associated with $\mathbf{X_{t-1}}$ (Lajaunie, 2021). In equation (1), γ_j and η_j are the vectors of coefficients for lagged values of y_{t-i} and π_{t-j} respectively, with p and q being the number of lags for the dichotomous and index variables. Kauppi and Saikkonen (2008) provided four particular specifications of the model (1): (1) a static model $\gamma = \eta = 0$; (2) a dynamic model with $\eta = 0$; (3) a dynamic model with $\gamma = 0$, and (4) a dynamic model with lagged binary and index variables, such that $\gamma, \eta \neq 0$.

The probability $\widehat{Pr}$ of a crisis occurring is calculated for each period under study, with the identification of a crisis determined by whether this probability exceeds a certain threshold. In this study, we employ three standard threshold criteria, each of which optimises the cut-off in a different way. As discussed in Hasse and Lajaunie (2021), the accuracy measure (AM) criterion aggregates the number of periods in which a crisis occurs and does not occur to give an optimal value that maximises the number of correctly identified periods. The credit-scoring approach (CSA) criterion looks at the sensitivity and specificity of correctly-identified crises: *sensitivity* gives the proportion of correctly identified crisis periods, whereas *specificity* is the proportion of correctly identified calm (non-crisis) periods, with the threshold value minimising the absolute difference between the two. The Noise-to-Signal Ratio (NSR) criterion (originally proposed by Kaminsky et al., 1998) represents the ratio of false alarms relative to correct alarms, with the threshold value minimising such relative errors. In general, the AM and CSA criteria are relatively similar and identify more crises than the NSR criterion, which typically has higher values and is thus more conservative. As such, AM and CSA are more likely to produce Type 1 errors (false positives), and NSR more Type 2 errors (false negatives).

Model Analysis For each model in Section 3, we evaluate its ability to successfully predict the binary response variable of interest. Based on the literature on signal detection, and juxtaposing possible outcomes against prediction in a so-called *confusion matrix*, in Table 1 we include key summary measures used for assessing the performance of each model. These measures for evaluating are used in our accuracy analysis through the receiver operating characteristic (ROC) curve and the area under the curve (AUC).

Table 1: Confusion Matrix Analysis

Method	Description	Formula
Accuracy	Correct identification	$(TP+TN)/(TP+TN+FP+FN)$
Precision	Positive predictive value	$TP/(TP+FP)$
Sensitivity	True-positive rate	$TP/(TP+FN)$
Specificity	True-negative rate	$TN/(FP+TN)$
1 − Specificity	False-positive rate	$FP/(FP+TN)$
Kappa	Degree of agreement	$(Accuracy-R)/(1-R)$
	where $R =$	$\frac{((TP+FN)(TP+FP)+(FP+TN)(FN+TN))}{(TP+TN+FP+FN)^2}$

The four possible outcomes are true positive (TP), false positive (FP), false negative (FN), or true negative (TN). Table adapted from Bowers and Zhou (2019). 1 − Specificity is used in the ROC curve. Kappa is Cohen's Kappa Score.

2.2 Data Sources for Modelling

An early warning model has two main non-latent empirical components: the binary variable y_t indicating a crisis and the vector of exogenous explanatory variables, $\mathbf{X_t}$. As the definition of a crisis requires a separate reflection, first we present the data sources for the explanatory variables, before defining what may constitute a migration crisis in Section 2.3. Asylum applications are accessed from Eurostat; Global event data from GDELT; Google searches from Google Trends; Ukrainian inflation from the State Statistics Service; US trade data from FRED St Louis; and exchange rates from the IMF.

The main variable of interest, underpinning the binary crisis response, is the number of asylum applications. We use first-time applications, as they are more likely to capture newly-arrived applicants. The relevant data are available at a monthly frequency from 1999:01 onwards in the Eurostat database.[3] The aggregation of the asylum applications for all these countries together yield the values used in our analysis.

A primary aim of this paper is to construct and test an early warning model for asylum applications with a six-month lead time. As such, the crisis indicators are taken a time t, while the explanatory variables are taken from six months previously $(t-6)$, so that e.g. a crisis in July 2015 is predicted using the explanatory variables up until January 2015. For each variable (where appropriate), we include the lag and difference

[3]The countries reporting to Eurostat for the period are the EU27 (as of 2007), Iceland and Norway. For the period 1999:01-2007:12 the table MIGR_ASYCTZM, and 2008:01 to present, table MIGR_ASYAPPCTZM is used. The reporting countries include 'EU+' countries: EU28 (with UK reporting until November 2020), Iceland, Norway, Switzerland, Lichtenstein and Montenegro.

for up to 12 months in the LASSO estimation which selects the explanatory variables. In terms of possible predictors, attempts to forecast civil unrest have been recently made, in particular with reference to the Arab Spring, using data from social media, and (protest) events using the GDELT [Global Database of Events, Language and Tone] data (Wu and Gerber, 2018).[4]

For Google Trends data, we looked at the relative frequencies of internet searches for migration-related terms, in the spirit of Böhme et al. (2020) and Avramescu and Wiśniowski (2021), adjusting them to more specific circumstances and using the sending country's language – Arabic for Syria, and Russian and Ukrainian for Ukraine.

2.3 What is a Migration Crisis?

A common feature shared by many, if not all early warning models is that their response is typically a binary variable identifying a *migration crisis*. This, however, sidesteps the question, how a crisis should be defined. In macroeconomics, for example, a conventional definition of recession is two consecutive quarters of negative growth. For migration, in particular asylum-related migration, this is much less clear. In existing work, Napierała et al. (2022) used migration exceeding one standard deviation above the historical average (or trend) for the distribution of asylum applications lodged in the EU as one example of such a definition, which turned out to be a rather sensitive threshold, as well as two standard deviations, as an upper limit. Approaches based on standard deviations is not always guaranteed to work, though. Our investigation focuses on the period 2008–2022, which contains a large variation in the number of asylum applications.

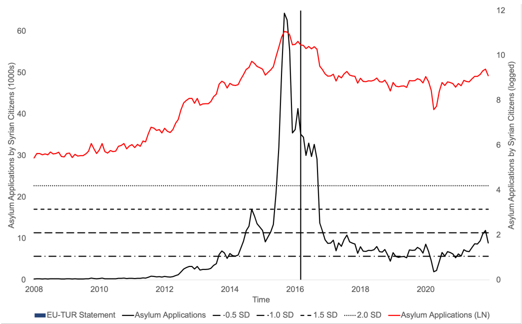
Figures 1a and 1b show the number of asylum applications and corresponding standard deviations (left axis), and the log-transformed number of applications (right axis), for Syria and Ukraine respectively.[5] For purpose of defining a crisis at present, applying the threshold of one standard deviation (for the period 2008–2021 inclusive), for the absolute

[4]At a general level, GDELT provides an invaluable source of information about events that occur in each country, categorising them using the extended codebook using the CAMEO format (for CAMEO codes, see website). There are many available aspects of GDELT data, particularly in GDELT 2.0, however for our purposes, we focus on the *events* database, which includes the numbers of events, average media tone of reporting, the so-called *Goldstein Scale*, summarising the severity of events and the tone, and a number of other ways to analyse the events.

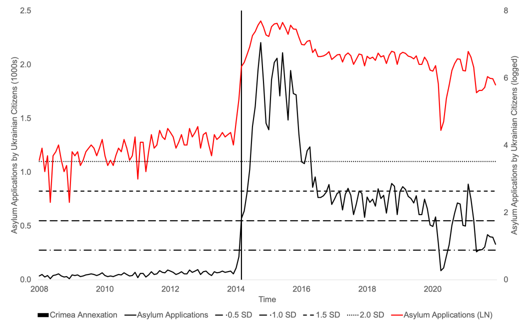
[5]We plot the log-transformed numbers in both figures due to the large change in 2015 and 2014, respectively, which (somewhat) mask the true scale of changes afterwards.

values would imply warnings issued before the respective 2015 peaks, and for Ukraine the average values would remain above the threshold post 2015.

From Figure 1a, we can see that for Syria, from December 2016, the monthly number of asylum application falls below one standard deviation, yet the absolute values still exceeded 75,000 in 2017, 2018, 2019 and 2021. Are these years a crisis? In one sense, yes, since there is a continued high inflow of asylum seekers. In another sense, no, since the numbers are smaller number than previously, especially at the 2015–16 peak. Defining past crises is equally important as defining a current crisis to make the model estimation more accurate.



(a) Syrian citizens



(b) Ukrainian citizens

Figure 1: Number of monthly asylum applications

The number of monthly asylum applications made by Syrian and Ukrainian citizens in EU+ (EU, EFTA, UK, ME) 2008:01–2021:12. The vertical axis shows the number of monthly asylum applications, in thousands. The fall in early 2020 is explained by the COVID-19 pandemic. Thresholds based on full-sample multiples of standard deviations (0.5, 1.0, 1.5 and 2.0SD) are shown by the horizontal lines. The vertical line on Figure 1a highlights the March 2016 statement between the EU and Turkey on Syrian asylum seekers (see: European Council) and on Figure 1b the annexation of Crimea (March 2014).

The Ukrainian case (Figure 1b) is somewhat different. From March 2014, except for the first COVID-19 lockdown and a handful of other months post the original COVID-19 outbreak, asylum applications exceed the full-period one standard deviation threshold.[6]

In addition, a lot of conflict-related displacement is internal: the recent estimates for Syria indicate 6.7 million internally displaced persons at the end of 2021, and for Ukraine nearly 7.0 million at the end of August 2022, the latter based on a survey carried out

[6]Since the Russian invasion in February 2022, the EU policy granted Ukrainian citizens rights to stay and work in the EU (initially for a year, with possible renewal for up to three years), by applying the Temporary Protection Directive based on citizenship and residence grounds, removing the need to individually claim asylum. The temporary protection status is easier to obtain, and comes with many rights (residence, work, choice of an EU country, access to services) but is also time-limited and potentially less stable than a refugee status, which is more difficult to secure. This legal change additional distorts indicators, such as the standard deviation, based on the whole period.

by the International Organization for Migration.[7] In particular, the displacement of a large number of Ukrainian citizens during 2014–15 occurred internally. The Ukrainian government reported that at the height of military operations in that period, 1.5 million people had been internally displaced (Jaroszewicz, 2019).

Two different, yet not mutually exclusive, questions are important here: (i) is there a significant stochastic movement (*change*) in the number of asylum applications, and (ii) is there a high volume of asylum applications. To address the former, we can use growth rates of the numbers of applications: they will be sensitive to changes at lower numbers, which are manageable from a policy perspective, but can hide crises if the numbers are too high but not growing compared to previous levels. No single indicator can unequivocally answer both questions, and a degree of subjectivity is inevitably required in defining a crisis. A static standard deviation-based approach would not be suitable to define a crisis, although a rolling 12-month standard deviation could offer an advancement on this. Due to the dynamics of applications for both countries, we see that for Syria a tighter definition is needed. Syria experiences large increases and exceptionally high flows over 4 years, while Ukraine has a much lower level and certain events rather than a full-scale civil war.

As such, our proposed binary response variable y_t has four components: the current period growth rate (g), with h denoting the number of months M over which the current period growth rate is calculated; a minimum growth rate that must have occurred in the previous twelve months (j), and whether the number of applications (x) in the current period exceeds a function f of the preceding 12-month rolling standard deviation SD, ($f(SD)$), and also exceeds a pre-set minimum value, X_{min}. If any one of the associated conditions, listed in the definition below, is not met, the value equals zero.

$$y_t = \begin{cases} 1 & g_{hM} \geqslant G\% \ , \ j \geqslant J\% \ , \ x \geqslant f(SD) \ , \ x \geqslant X_{min} \quad (2) \\ 0 & g_{hM} < G\% \mid j < J\% \mid x < f(SD) \mid \ x < X_{min} \quad (3) \end{cases}$$

[7]Source: IDMC - Internal Displacement Monitoring Centre, Syria and Ukraine country profiles, accessible via https://www.internal-displacement.org (as of 25 September 2022)

3 Model Specification

In this section, we present background to each case study, and describe the models. Contemporaneous models and six-month forecasts are used to assess the robustness using such models to forecast crises.

3.1 Syria

The first case study designed to test our model is the related to asylum applications of Syrian nationals (or people claiming Syrian citizenship) in the EU and neighbouring countries, with focus on the events of 2015–16. The Arab Spring protests led to political unrest across the region and to a full-scale civil war in Syria. Since then, increasingly more asylum seekers, especially Syrians, were trying to reach Europe. With some family members having left earlier on in this period, family reunification followed suit.

3.1.1 The Syrian crisis in the light of data

GDELT data In an attempt to identify possible explanatory variables that could serve as early warnings, first, we examine the number of protests and total number of political events in Syria in the period 2007:01-2022:04.[8] The number of protest peaks in 2011–12, and once the civil war starts, there is a further increase in the number of all events from 2012. During the Arab Spring, the protest intensity (number of protests divided by the number of total recorded events) was at its maximum in April 2011.[9]

Even though the number of protest events can tell a profound story, another important variable available in the GDELT database is the average *tone* of the reporting of the event in the media. The average tone, defined by the GDELT collection, is how positively or negatively the reports of the event are. The scores range from –100 (extremely negative) to +100 (extremely positive), though more commonly they range between –10 to +10, with 0 being considered neutral. We cannot identify the actual events or reports using this approach of GDELT, however, through a combination of actors, date and location we are able to isolate them and approximate the likely events. The average tone is useful

[8]GDELT data are unavailable for extraction for 23-25 January 2014 and 19 March 2014.

[9]Even though the fall in the number of protests does not perfectly coincide with the increase in asylum applications, there is a clear relationship between the events of the Arab Spring and the start of the increase in asylum applications.

in determining the context and seriousness of protests or conflict. Small riots are likely to receive a small negative score, whilst large negative scores would indicate more serious events.

For Syria, there are three notable changes in the tone of reporting: in early months of 2011, at the end of 2012 and, most pronounced, in winter 2014/15. There is a clear time lag between the changes in the average tone of reporting and the number of asylum applications. This information can help build our model, since the presence of any lags is crucial in the context of early warnings.

Macroeconomic data There are key parameters of macro-financial stability including exchange rates and inflation.[10] However, for some less developed countries, and in particular, conflict strewn or under economic sanctions from e.g. western economies, data are often unavailable or unreliable. A key example of data being reduced in quality (and thus its information value) is the official exchange rate: an example for the Syrian Pound (SYP) exchange rate against the USD. The series of data available from the International Monetary Fund (IMF), and other financial data outlets, stops towards the end of 2018. From a macroeconomic perspective, we can also see very clear effects of the Syrian civil war by looking at the US trade with Syria. This data are available to 2022, with some reporting lag to be expected. Sanctions placed on Syria by the US Government caused trade flows to collapse.[11] Exports peaked in September 2010 at $89.6 million, a year later they were down to just $3.4 million. Imports peaked in June 2011 at $116 million, falling to $1.9 million only six months later.

Google Trends There have already been some attempts to employ Google Trends data in detecting signal of increasing asylum applications. In Table 2 we present the variables we have prioritized for modelling. We included searchers on various border towns to the south and to the west of Syria, and Lebanon, and on some of the largest

[10] High rates of inflation, particularly hyperinflation, can be a push factor for would-be asylum applicants. High rates of inflation weakens the currency and can make it hard to access essential supplies. There are a number of examples of this, both in conflict or war-torn countries but also in politically-unstable countries such as Venezuela, which has seen over 3 million refugee since 2015 (Source: UNCHR Global Trends 2018, accessed 1 September 2022.) The high rates of inflation can be thus an early indication of an increase of asylum applications.

[11] In particular, Executive Order 13582 signed on 18th August 2011. Accessed 2 May 2022.

refugee camps. Regarding those who decide to reach Turkey, Google searches for border towns and refugee camps in Turkey did not return meaningful results, or the signal was too weak to show any patterns. The impact of the language varies, though to differing

Table 2: Google Trends

Google Trends search terms		
Airport* (A)	Goodbye* (A)	Reunification* (A)
Asylum* (A)	Greece (A)	Smuggler (A)
Diaspora (A)	Immigration* (A)	Smuggling* (A)
Emigration (A)	Lebanon (A)	Tell Abyad (A)
German Embassy* (E)	Mafraq, Jordan (A)	Tell Abyad (E)
Germany (A)	Refugee (A)	Turkey Border (A)

A list of Google Trends terms included in the exploratory analysis for the EWS model. (A) identified a search term in Arabic, whilst (E) denotes searches in English. The differences in language searches vary, with some being quite similar whilst others not. NB: Mafraq is a border town in Jordan. Data exported 9 March 2022. Variables with an asterisk are used in the models.

extent for different terms. The term 'reunification' increases in significance from August 2014 onwards, reaching a maximum in September 2015. This is in contrast to 'Goodbye', which peaks in October 2012, before reducing significantly. There is a similar pattern for 'immigration' and 'asylum', with 'immigration' returning higher search values before 2015, before 'asylum' started to dominate in 2014–16.

3.1.2 Early Warning System Models

Following the argumentation presented in Section 2.3, we define the dichotomous warning variables for the Syrian case study as a combination of four criteria: a period growth rate exceeding 50% over 6 or 12 months; at least one growth rate of greater than 50 or 100% in the previous 12 months, the number of asylum applications exceeding 2 standard deviations from the preceding 12 months, and over 300 asylum applications per month.

Explanatory variables include: data from Google Trends (labelled GT), where necessary, with superscript A indicating search terms in Arabic, otherwise these are topics or search terms in English. The GDELT reporting uses events in Syria, with sorting by event root code (ERC) with analysis of counts (Ct) or average tone (AT). US imports from Syria are denoted by $USImpSyr$, exports (Exp) and net exports (NX) to Syria follow equivalent notation. The lags for p and q, when no restrictions apply, as described in equation (1), and equal 1 unless otherwise stated. The matrices of explanatory variables, which are determined by the LASSO selection in R (Friedman et al., 2022), primarily consist of asylum applications, Google Trends data on different search terms, and GDELT data for the ERC codes 14 (protests), 18 (assault) and 19 (fight). Lags in explanatory variables denoted by L and first differences by Δ.

EWS at time T

The first model, for contemporaneous (at time T) warnings (model T-M1), is shown in equation (4). The binary response variable depicts 61 crisis periods, with the first one in August 2011. There are no restrictions on the lagged latent or dichotomous variables, $\gamma_j, \eta_j, \neq 0$ with lags of $p = q = 1$.

$$y_t = \begin{cases} 1 & \text{if} \;\; g_{12M} \geqslant 50\%, \;\; j \geqslant 100\%, \;\; x \geqslant 2SD \;, \;\; x \geqslant 300 \\ 0 & \text{if} \;\; g_{12M} < 50\% \mid j < 100\% \mid x < 2SD \mid x < 300. \end{cases}$$

$$\mathbf{X_t} = \big[\mathrm{SYRAsyApp}_{\Delta 12}, \mathrm{GTAsylum}^{\mathrm{A}}_{\mathrm{t}}, \mathrm{GTGoodbye}^{\mathrm{A}}_{\mathrm{t,L12}}, \mathrm{GTImmigration}^{\mathrm{A}}_{\Delta 11, \Delta 12}, \mathrm{GTReunification}^{\mathrm{A}}_{\Delta 11}, \mathrm{ERC14Ct}_{\mathrm{L8,L9}}\big] \quad (4)$$

The second contemporaneous model (T-M2), uses a minimum requirement of 50% six-month growth rate as described in equation (5), identifying 26 crisis periods, starting from August 2012. There is a lag of $p = 1$, with a restriction on the latent variable, $\eta_j = 0$.

$$y_t = \begin{cases} 1 & \text{if} \;\; g_{6M} \geqslant 50\%, \;\; j \geqslant 100\%, \;\; x \geqslant 2SD \;, \;\; x \geqslant 300 \\ 0 & \text{if} \;\; g_{6M} < 50\% \mid j < 100\% \mid x < 2SD \mid x < 300. \end{cases}$$

$$\mathbf{X_t} = \big[\mathrm{SYRAsyApp}_{\Delta 5}, \mathrm{GTAsylum}^{\mathrm{A}}_{\mathrm{t,L1}}, \mathrm{GTGoodbye}^{\mathrm{A}}_{\mathrm{L3}}, \mathrm{GTImmigration}^{\mathrm{A}}_{\Delta 7,\mathrm{L1}}\big] \quad (5)$$

The third contemporaneous model (T-M3) is defined in equation (6), reducing the requirement for the minimum growth rate from the preceding 12 months, j, to 50%, with the binary response variable identifying 64 crisis periods, starting from June 2011. There is a restriction of $\gamma_j = 0$ on the dichotomous variable, and a lag of $q = 3$.

$$y_t = \begin{cases} 1 & \text{if} \;\; g_{12M} \geqslant 50\%, \;\; j \geqslant 50\%, \;\; x \geqslant 2SD \;, \;\; x \geqslant 300 \\ 0 & \text{if} \;\; g_{12M} < 50\% \mid j < 50\% \mid x < 2SD \mid x < 300. \end{cases}$$

$$\mathbf{X_t} = \big[\mathrm{SYRAsyApp}_{\Delta 12}, \mathrm{GTAsylum}^{\mathrm{A}}_{\mathrm{t}}, \mathrm{GTGoodbye}^{\mathrm{A}}_{\mathrm{t,L2}}, \mathrm{GTReunification}^{\mathrm{A}}_{\Delta 11}, \\ \mathrm{ERC14Ct}_{\mathrm{L2,L9}}, \mathrm{ERC19AT}_{\Delta 9}, \mathrm{ERC19Ct}_{\Delta 12},\big] \quad (6)$$

EWS at a six-month horizon

The following models use the data with a six-month lag, with binary response variable based on the asylum applications at time t, whereas the explanatory variables in the matrix $\mathbf{X_t}$ are relative to the data from 6 months before $(t-6)$. The explanatory variables are defined as before. The binary variables and corresponding matrices for models in the second panel of Table 4 are shown in equations (7) to (9).

The first model with a six-month warning horizon (T6-M1), is defined in equation (7), where there are no restrictions $\gamma_j, \eta_j \neq 0$ and $p = q = 2$. The crisis indicator is the same as described in equation (4).

$$y_t = \begin{cases} 1 & \text{if } \ g_{12M} \geqslant 50\%, \ j \geqslant 100\%, \ x \geqslant 2SD \ , \ x \geqslant 300 \\ 0 & \text{if } \ g_{12M} < 50\% \mid j < 100\% \mid x < 2SD \mid x < 300. \end{cases}$$

$$\mathbf{X_t} = \big[\mathrm{SYRAsyApp}_{\mathrm{L12}}, \mathrm{GTAirport}_{\mathrm{L1,L3}}, \mathrm{GTAsylum}^{\mathrm{A}}_{\Delta 11}, \mathrm{GTGerEmb}_{\mathrm{L2}}, \mathrm{GTGoodbye}^{\mathrm{A}}_{\mathrm{t,L1,L2,L3,L7,L8}}, \\ \mathrm{GTReunification}_{\Delta 11,\mathrm{L9,L10}}, \mathrm{ERC14Ct}_{\mathrm{L2,L3,L5}}\mathrm{ERC18AT}_{\mathrm{L12}}, \mathrm{ERC18Ct}_{\Delta 11,\Delta 12}, \mathrm{ERC19Ct}_{\Delta 12}\big] \quad (7)$$

Equation (8) describes the second model with a six-month horizon (T6-M2), corresponding to (6), with a restriction of $\gamma_j = 0$ and a lag of $q = 1$.

$$y_t = \begin{cases} 1 & \text{if } \ g_{12M} \geqslant 50\%, \ j \geqslant 50\%, \ x \geqslant 2SD \ , \ x \geqslant 300 \\ 0 & \text{if } \ g_{12M} < 50\% \mid j < 50\% \mid x < 2SD \mid x < 300. \end{cases}$$

$$\mathbf{X_t} = \big[\mathrm{SYRAsyApp}_{\mathrm{L12}}, \mathrm{GTAirport}^{\mathrm{A}}_{\mathrm{L3}}, \mathrm{GTGerEmb}_{\mathrm{L2}}, \mathrm{GTGoodbye}^{\mathrm{A}}_{\mathrm{t,L1,L2,L3,L7}}, \\ \mathrm{GTReunification}^{\mathrm{A}}_{\Delta 11,\mathrm{L9,L10}}\mathrm{GTSmuggling}^{\mathrm{A}}_{\mathrm{L3}}, \mathrm{ERC14Ct}_{\mathrm{L2,L3,L4}}, \mathrm{ERC18Ct}_{\Delta 11}, \mathrm{ERC18AT}_{\mathrm{L12}}\big] \quad (8)$$

The final model (T6-M3), based on six-month growth rates, is presented in equation (9), with no restrictions such that $\gamma_j, \eta_j \neq 0$ and a lag of $p = q = 1$, identifying 34 crisis periods, starting from July 2011.

$$y_t = \begin{cases} 1 & \text{if } \ g_{6M} \geqslant 50\%, \ j \geqslant 50\%, \ x \geqslant 2SD \ , \ x \geqslant 300 \\ 0 & \text{if } \ g_{6M} < 50\% \mid j < 50\% \mid x < 2SD \mid x < 300. \end{cases}$$

$$\mathbf{X_t} = \big[\mathrm{USImpSyr}_{\Delta 9}, \mathrm{USNXSyr}_{\mathrm{L8}}, \mathrm{GTGoodbye}^{\mathrm{A}}_{\mathrm{t,L9}}, \mathrm{ERC18AT}_{\Delta 3}\big] \quad (9)$$

3.2 Ukraine

The Ukraine case study focuses on two periods: the build-up of the crisis since 2013–14, starting with the pro-European *Euromaidan* protests (Zelinska, 2017) and the ensuing Russian illegal annexation of Crimea, and the run-up to February 2022, with mass-scale population displacement caused by the Russian invasion. In the intervening period, there has been intermittent fighting and low-intensity conflict in the eastern Ukrainian regions of Donetsk and Luhansk. In 2022, in the first four months after the full-scale Russian invasion, the UNHCR estimated that some 7.5 million Ukrainians crossed the borders into neighbouring countries, of whom around 4 million entered into Poland alone – while in the same period 2.5 million have crossed the border back into Ukraine.[12]

3.2.1 The Ukraine crisis in the light of data

GDELT data An important caveat of GDELT is that the indicators that might have worked for Syria, do not necessarily have to work as well for Ukraine. For Ukraine, analysing protests and number of other political events *separately* is particularly important. There is a spike in the relative protest *intensity* for the 2013–14 crisis, but the 2021–22 crisis does not have the same spike in protest intensity due to the protests being dwarfed by the number of other political events. Similarly to the Syrian example, an alternative approach relies on looking at the *average tone* of reporting of all political events. Even with reasonable lags, there does not appear to be a significant drop in tone to coincide with the events that trigger an increase in asylum applications represented in this data series as with Syria.

Macroeconomic data There are a number of possible *leading* macroeconomic indicators that can signal upcoming crises weeks or months ahead, in particular, exchange rates. Since the Ukrainian Hryvnia had a peg to the USD until inflationary pressures caused it to become a managed float in 2015, an alternative to the nominal exchange rates, we include in our analysis the real effective exchange rate (REER) based on the consumer price index (CPI), which is not affected by pegging or interest rates. The real effective exchange rate is an index, is defined by the IMF as "...a measure of the value

[12]Source: UNHCR, accessed on 15 June 2022.

of a currency against a weighted average of several foreign currencies) divided by a price deflator or index of costs". An increase in the REER thus indicates a decline in trade competitiveness.

As for the trade balance, the decline in trade concerning Ukraine was not externally imposed, as was the case with the US ban on trade with Syria, but rather the political (and consequentially economic) situation in Ukraine has dictated the change in trade flows. The US has been shifting between a trade surplus and trade deficit with Ukraine, with a general trend towards a trade surplus. Following the start of both phases of the conflict, there was a general decline in trade both for imports and exports. At the same time, for Russia, recent figures from the US Bureau of Economic Analysis show a significant decline in trade since the invasion of Ukraine, due in part to the sanctions imposed by the West. There was also a decrease in trade flows after the 2014 annexation of Crimea but not to the same extent. The US has consistently run a large trade deficit with Russia. In summary, the changes in trade flows can indicate issues or conflicts that arise within an economy, or in this case, are imposed by external forces.

Google Trends Table 3 shows a selection of the search terms that were included in the initial selection. The searches included terms in Russian and Ukrainian, with border towns in Ukraine, as well as Kraków, as one main Polish destination city. The terms included in the analysis exhibited patterns which could help predict an increase in asylum applications, but the ones selected for further analysis were those selected in the LASSO modelling process.

3.2.2 Early Warning Models

As in the Syrian case study, we first introduce the binary response and explanatory variables. All response variables in this case study require a value exceeding one or two standard deviations, with a period growth rate of 10 or 25% over the previous six or 12 months, and at least one growth rate of greater than 25% in the previous 12 months, with a minimum value of 50 asylum applications per month. The lags for p and q, when no restrictions apply, as described in equation (1), are equal 1 unless otherwise stated.

In addition to the data definitions introduced before, for Google Trends data a su-

Table 3: Google Trends

Google Trends search terms			
Airport (R)	Border (R)	Krakow* (P)	Russia* (U)
Airport (U)	Diaspora	Mali Selmentsi (U)	Shehyni (U)
Asylum* (R)	Duolingo	Moldova*	Slovakia
Asylum* (U)	European Union*	Passport	Train Station
BBC*	Europe (R)	Poland (R)	US Dollar
BBC (ST)	Europe* (U)	Poland (U)	Uzhhorod (U)
Bitcoin	Germany	Zloty	
Border (R)	Goodbye (U)	Russia* (R)	

A list of Google Trends terms included in the exploratory analysis for the EWS model. (R) identified a search term in Russian, whilst (U) denotes searches in Ukrainian, and (P) in Polish. (ST) specifies a search topic rather than a search term. Data exported 6 November 2022. Variables with asterisks were selected by LASSO and were used in the models.

perscript R indicates Russian language, and U stands Ukrainian. Additional GDELT variables include the number of mentions ($NoMe$), Goldstein Scale (GS), count of negative GS events ($NegGSCt$), and protest intensity ($ProtestInt$). Actors include $RM =$ Russian Military, $RGM =$ Russian Government and Military, with event root codes given in the superscript.[13] For trade to/from the US, $USImpUKR$, denotes US imports from Ukraine, exports (Exp) and net exports (NX) to Ukraine and trade flows with Russia follow similar notation. The real exchange rates for Russia and Ukraine are given by $RusRealXR$ and $UkrRealXR$ respectively, and Ukrainian inflation by $UKRInfl$.

EWS at time T

The first two contemporaneous models are shown in equations (10) (T-M1) and (11) (T-M2), respectively. The explanatory variables, $\mathbf{X_t}$ include the difference of asylum applications by Ukrainian citizens over a 12-month period, US exports and imports to/from Russia at varying lag lengths, and Google Trends searches for Asylum in Russian at varying lag lengths, as well as the number of protests in Ukraine for the previous month. The binary response variable y_t given in equation (10) identified 36 crisis months, with the first one in July 2013. There are restrictions on the lagged latent or dichotomous variables, $\gamma_j, \eta_j = 0$, with $p = q = 1$.

[13] Numbers 0120 indicate all codes, with 1020 for codes 10-20.

$$y_t = \begin{cases} 1 & \text{if } \ g_{12M} \geqslant 20\%, \ j \geqslant 25\%, \ x \geqslant 2SD \ , \ x \geqslant 50 \\ 0 & \text{if } \ g_{12M} < 20\% \mid j < 25\% \mid x < 2SD \mid x < 50. \end{cases}$$

$$\mathbf{X_t} = \left[\mathrm{UKRAsyApp}_{\Delta 12}, \mathrm{USExpRUS}_{\mathrm{L8,L9}}, \mathrm{USImpRUS}_{\mathrm{L12}}, \mathrm{GTAsylum}^{\mathrm{R}}_{\mathrm{t,L2,L6}}, \mathrm{NoProtestsUKR}_{\mathrm{L1}}\right] \quad (10)$$

The binary variable in equation (11), reducing the growth rate threshold to 10% and SD to one, indicates 46 crises, with the first two occurring in June and July 2013. There are restrictions on the latent variable, $\gamma_j = 0$, with a lag $q = 1$.

$$y_t = \begin{cases} 1 & \text{if } \ g_{12M} \geqslant 10\%, \ j \geqslant 25\%, \ x \geqslant 1SD \ , \ x \geqslant 50 \\ 0 & \text{if } \ g_{12M} < 10\% \mid j < 25\% \mid x < 1SD \mid x < 50. \end{cases}$$

$$\mathbf{X_t} = \left[\mathrm{UKRAsyApp}_{\Delta 12}, \mathrm{USExpRUS}_{\mathrm{L3}}, \mathrm{USImpRUS}_{\mathrm{L9}}, \mathrm{GTAsylum}^{\mathrm{R}}_{\mathrm{t,L2,L10}}\right] \quad (11)$$

EWS at a six-month horizon

The models with a six-month horizoninclude a wide range of Google Trends, macroeconomic (trade and exchange rates) and GDELT data. The first six-month horizon model (T6-M1) in equation (12) uses the same binary response variable as (10). There are no restrictions on the latent of dichotomous variables, $\gamma_j, \eta_j \neq 0$, with a lag $p, q = 2$.

$$y_t = \begin{cases} 1 & \text{if } \ g_{12M} \geqslant 20\%, \ j \geqslant 25\%, \ x \geqslant 2SD \ , \ x \geqslant 50 \\ 0 & \text{if } \ g_{12M} < 20\% \mid j < 25\% \mid x < 2SD \mid x < 50. \end{cases}$$

$$\mathbf{X_t} = \left[\mathrm{UKRAsyApp}_{\Delta 7}, \mathrm{USExpRus}_{\mathrm{L1,L2,L3}}, \mathrm{USImpRus}_{\mathrm{L4,L6}}, \mathrm{RusRealXR}_{\mathrm{L4}}, \mathrm{GTAsylum}^{\mathrm{R}}_{\mathrm{t,L4,L10,\Delta 12}}, \mathrm{GTAsylum}^{\mathrm{U}}_{\mathrm{L8}}, \mathrm{GTEurope}^{\mathrm{U}}_{\Delta 11}, \mathrm{GTEU}_{\Delta 11}, \mathrm{GTRussia}^{\mathrm{U}}_{\mathrm{L7}}, \mathrm{GTKrakowPol}_{\mathrm{L9}}, \mathrm{ProtestUKRCtLn}_{\Delta 12}, \mathrm{ProtestIntUKR}_{\mathrm{L9}}\right] \quad (12)$$

The second model (T6-M2), defined in (13), uses that same binary variable as in (11). There are no restrictions on the latent or dichotomous variables, with $p = q = 1$.

$$y_t = \begin{cases} 1 & \text{if } \ g_{12M} \geqslant 10\%, \ j \geqslant 25\%, \ x \geqslant 1SD \ , \ x \geqslant 50 \\ 0 & \text{if } \ g_{12M} < 10\% \mid j < 25\% \mid x < 1SD \mid x < 50. \end{cases}$$

$$\mathbf{X_t} = \left[\mathrm{UKRAsyApp}_{\Delta 7}, \mathrm{USExpRus}_{\mathrm{L1,L2,L3}}, \mathrm{USImpRus}_{\mathrm{L3}}, \mathrm{RusRealXR}_{\mathrm{L5,L10}}, \mathrm{GTAsylum}^{\mathrm{R}}_{\mathrm{t,L4}}, \mathrm{GTRussia}^{\mathrm{R}}_{\Delta 11}, \mathrm{ProtestIntUKR}_{\mathrm{L8}}\right] \quad (13)$$

The third model (T6-M3) with a six-month horizon uses a six-month growth rate with a current period growth rate requirement of 10%, identifying 51 crisis periods. There are no restrictions on the latent and binary variables, with lags $p = q = 3$.

$$y_t = \begin{cases} 1 & \text{if } g_{6M} \geqslant 10\%,\ j \geqslant 25\%,\ x \geqslant 2SD\ ,\ x \geqslant 50 \\ 0 & \text{if } g_{6M} < 10\% \mid j < 25\% \mid x < 2SD \mid x < 50. \end{cases}$$

$$\begin{gathered}\mathbf{X_t} = \big[\mathrm{UKRAsyApp_t}, \mathrm{A1UKRA2RUSNoMe}^{\mathrm{ERC0120}}_{\mathrm{L4}}, \mathrm{A1RGMA2UKRGS}^{\mathrm{ERC0120}}_{\Delta 9,\mathrm{L9}}, \mathrm{PutinCt}^{\mathrm{ERC0120}}_{\Delta 1}, \\ \mathrm{RMNoMe}^{\mathrm{ERC1020}}_{\Delta 7}, \mathrm{RMAT}^{\mathrm{ERC1020}}_{\Delta 8}, \mathrm{RMCt}^{\mathrm{ERC0120}}_{\mathrm{L7}}, \mathrm{USExpRus_{L3}}, \mathrm{UKRInfla_{L12}}, \mathrm{RMNegGSCt_{L8}}, \\ \mathrm{GTAirport}^{\mathrm{R}}_{\mathrm{L11}}, \mathrm{GTAsylum}^{\mathrm{R}}_{\mathrm{L1}}, \mathrm{GTAsylum}^{\mathrm{U}}_{\Delta 4,\Delta 8}, \mathrm{GTBitcoin_{L4}}, \mathrm{GTBorder}^{\mathrm{U}}_{\Delta 8}, \mathrm{GTEU}_{\Delta 12}, \mathrm{GTEurope}^{\mathrm{U}}_{\Delta 12}, \\ \mathrm{GTGermany}_{\mathrm{t},\Delta 6}, \mathrm{GTGoodbye}_{\Delta 7,\Delta 11}, \mathrm{GTKrakowPol_t}, \mathrm{GTMaliSelmentsi_{L11}}, \\ \mathrm{GTMoldova}_{\Delta 11,\mathrm{L5}}, \mathrm{GTRussia}^{\mathrm{U}}_{\mathrm{L1,L2}}, \mathrm{GTSlovakia_{L12}}, \mathrm{GTTrainStation}_{\Delta 12}, \mathrm{AllEventsUKRCt}_{\Delta 7}\big] \end{gathered} \tag{14}$$

4 Results

In this section we present the results for each of the case studies presented in Section 3.

4.1 Syria

The first part of the results examines the results from the models in Section 3.1.2, with the second part evaluating model performance.

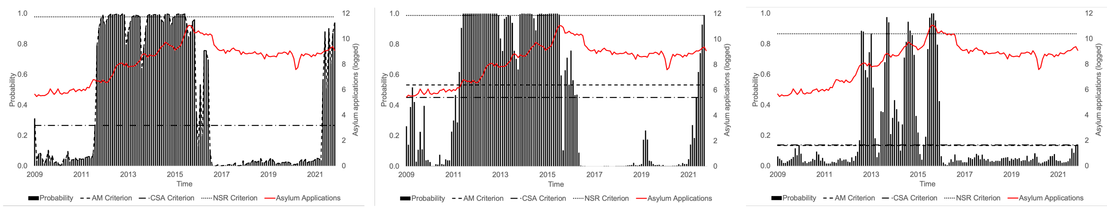
4.1.1 Early Warning Model Results

In this section, we report the results of the EWS modelling with respect to the probability of identifying crisis periods, and use the confusion matrix results to statistically analyse the model accuracy using the criteria described in Table 1.

The results for the three contemporaneous models are shown in Figure 2. The first plot, Figure 2a, shows the high probabilities of reaching the tight NSR criterion with the first warning for January 2012, and crises identified continuously from September 2011 for the AM, and October 2011 for the CSA criterion until June and April 2016 respectively. There are two false negatives using the AM criterion August 2011 and May 2021. With four of the eight false positives occurring between December 2015 - July 2016, which didn't meet the definition of a crisis in equation (4). The false negatives in May 2021 could be misleading due to the COVID lockdowns of 2020–21. Probabilities close to the NSR criterion indicate a more severe crisis, whilst those barely exceeding AM or CSA denote lower severity.

Figure 2b has the same threshold for growth rates and asylum applications, but the growth rate is calculated over a six-month period, as in (5). The first crisis post-Arab

Spring is identified in August 2012, nearly a year later than for the 12-month growth rates, at which time the asylum applications are around 2,000 per month and rising quickly. The only false negative occurs in October 2021, however a series of false positives occur during the early months of 2013 and 2014, and in the winters of 2014/15 and 2015/16, indicating that a potential issue may arise from seasonality, as the number of applications remain high. The third model uses 12-month growth rates but with a lower threshold, as per (6), with the results shown in Figure 2c. The first warning is issued in July 2011, following a false negative in June 2011, and the probability continues to exceed the AM threshold until November 2015, (false positive is identified for August 2013), with a mixture of outcomes in the varying period of November 2015 to July 2016, and most recently in October to December 2021 following false negatives in April, May, June and August 2021.

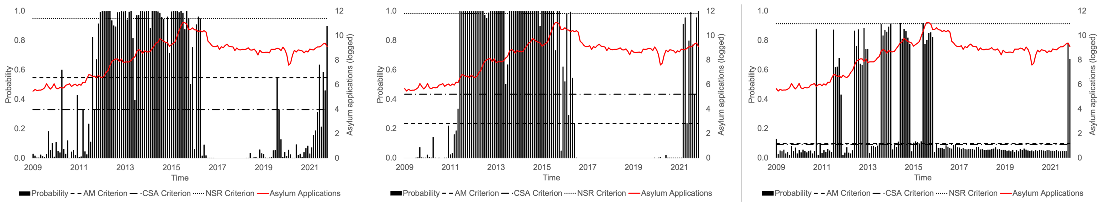


(a) Model T-M1 (b) Model T-M2 (c) Model T-M3

Figure 2: EWS results from Syria for the contemporaneous model

EWS results for Syria using contemporaneous data. The horizontal lines identify the AM, CSA, and NSR criteria. The vertical bars show the probability at the current period of a crisis occurring with the red line being the natural log of asylum applications. For Model T-M1, the threshold values are: AM=0.1043, CSA=0.2660, NSR=0.9780; Model T-M2: AM=0.1342, CSA=0.1378, NSR=0.8670; Model T-M3: AM=0.5336, CSA=0.4508, and NSR=0.9890. Source: Eurostat (asylum data); own elaboration from EWS results.

Figure 3 shows the corresponding three models for the six-month horizon. The probability and criteria from the first model, as defined in equation (7), are shown in Figure 3a. The first alerts of a crisis is for October 2011, which, as the data is at a six-month offset, would indicate a warning is given in April 2011, with a false negatives given for August, September and November. The crises are correctly assessed until a false positive in December 2015, followed by true negatives are experienced until a false negatives in May, June and November 2021. The second model with the six-month offset, described in (8), yields probabilities and threshold values shown in Figure 3b. The first warnings are given for August 2011, so identified in February 2011, with false positives in August



(a) Model T6-M1 (b) Model T6-M2 (c) Model T6-M3

Figure 3: EWS results from Syria at a six-month horizon

EWS results for Syria at a six-month horizon. The horizontal lines identify the AM, CSA, and NSR criteria. The vertical bars show the probability at the current period of a crisis occurring with the red line being the natural log of asylum applications. For Model T6-M1, the threshold values are: AM=0.5470, CSA=0.3296, NSR=0.9498; Model T6-M2: AM=0.2354, CSA=0.4339, NSR=0.9824; Model T6-M3: AM=0.0983, CSA=0.0915, and NSR=0.9126. Source: Eurostat (asylum data); own elaboration from EWS results.

2013, January and March 2016. True values are then observed until July 2021 with a false positive, again in September 2021.

The final model with the six-month offset is described in equation (9), with probabilities and criteria shown in Figure 3c. As a consequence of using a six-month growth rate to define a crisis, fewer crisis periods are identified, and as shown in Table 4, the model performs worse overall. The first crisis signal of July 2011 is a false negative based on all three criteria, with other false positives in early 2011. The model improves from 2012 onwards, though some false outcomes occur.

4.1.2 Confusion matrix and model performance

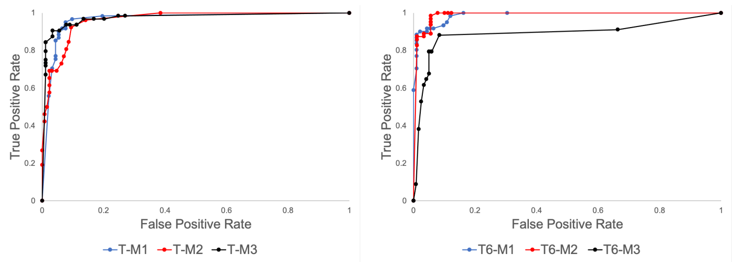
The results of classification are presented in a *confusion matrix*, in Table 4. We provide the results for using the threshold maximising the measures, usually the AM criterion; since NSR is likely to provide a larger number of false negatives, while CSA provides more false positives. In addition, an ROC curve plots true positive rate (TPR) vs the false positive rate (FPR), or sensitivity vs 1–specificity, at different classification thresholds, set at probability intervals of 0.05. The ROC curves shown in Figure 4 graphically demonstrate that all of the models are an improvement on a 50-50 chance, as the lines are all above the 45-degree line.

Depending on the strand of literature, different criteria may be preferable, for example Cohen's Kappa, AUC or MSE (see, for example, Ben-David, 2008). For our analysis, the results shown in the last three columns of Table 4 give consistent ranking of T6-M2, T6-

Table 4: Confusion matrix and contingency proportions

Model	TP	FP	FN	TN	Acc.	Prec.	Sens.	Spec.	Kappa	AUC	MSE
Signal detection characteristics for the contemporaneous models											
T-M1	59	8	2	86	93.55%	0.8806	0.9672	0.9149	0.8671	0.9641	0.0603
T-M2	24	12	2	117	90.96%	0.6667	0.9231	0.9070	0.7196	0.9620	0.0588
T-M3	58	3	6	86	94.12%	0.9508	0.9063	0.9663	0.8783	0.9732	0.0530
Signal detection characteristics at the six-month horizon											
T6-M1	55	2	6	91	94.81%	0.9649	0.9016	0.9785	0.8902	0.9873	0.0451
T6-M2	64	5	0	86	96.77%	0.9276	1.0000	0.9451	0.9342	0.9880	0.0375
T6-M3	30	11	4	110	90.32%	0.7317	0.8824	0.9091	0.7369	0.8911	0.0758

The confusion matrix details the accuracy of the results by comparing the predicted outcome with the actual outcome. The four outcomes are true positive (TP), false positive (FP), false negative (FN), or true negative (TN). The sum of outcomes differs depending on binary variable selected. Results for the models presented below using the AM criterion. The models are evaluated with Accuracy (Acc), Precision (Prec), Sensitivity (Sens), Specificity (Spec), Kappa, and Area Under the Curve (AUC) and Mean Squared Error (MSE).



(a) Contemporaneous (b) 6 Month Horizon

Figure 4: ROC Curve for Syria Models

The ROC plots for the six models analysed in Table 4. The horizontal axis is the False Positive Rate (FPR) and the vertical axis the True Positive Rate (TPR). For each model, the blue line corresponds to model 1 in the respective panel of Table 4, red for model 2, and black for model 3.

M1, T-M3, and T-M2 with the two poorer performing models differing with the Kappa measure. In terms of minimising the type I (FPR, 1–Specificity) and type II (FNR, 1–Sensitivity) errors, the best results are obtained for models T6-M1 (0.9785), and T6-M2 (1.0000) respectively, with model T6-M3 performing worst in terms of sensitivity, and T-M2 for specificity - notably both using six-month growth rates.

4.1.3 Discussion

From a policy perspective, the results of a six-month horizon model outperforming and providing an efficient indicator of a forthcoming crisis are promising, and potentially helping with preparedness. However, in the light of these findings, a larger question looms: is there another migration crisis coming? In each set of results, we can see periods of concern, indicated by higher probabilities, preceding the large increase in migration flows. In the increases in asylum applications towards the end of 2021, showed levels not seen since 2016. Of course, one consequence of the COVID-19 pandemic is that some asylum applications are delayed. However, the alerts given by other indicators suggest that at least some of this increase may be sustained. Still, in comparison with the magnitude of the Ukraine 2022 crisis, these increases in the numbers of new Syrian asylum applications would be much easier to manage, even if the legal treatment of these two groups of asylum seekers is different, as discussed before. Besides, some of the models featured in this section may suffer from issues identified elsewhere in the literature (e.g. Filippopoulou et al., 2020), such as the inclusion of post-crisis data, which additionally strengthens the case for using models using time horizons of at least a few months ahead.

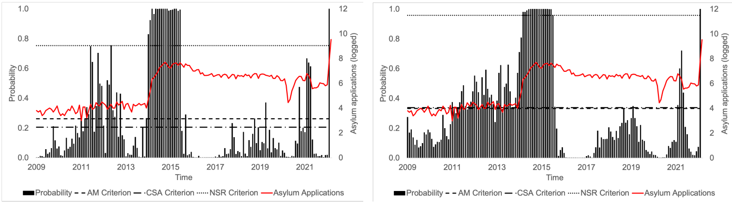
4.2 Ukraine

The first part of the results examines the results from the models in Section 3.2.2, with the second part evaluating model performance.

4.2.1 Early Warning Model Results

For Ukraine, Figure 5 shows the probability and corresponding thresholds for the contemporaneous models, which were unable to predict the crisis in February 2022 (only March). This is more likely to reflect poorer selection of explanatory variables for these models. In Figure 5a, the first crisis warning is given in January 2014. The true positives continue until July 2015, before becoming true negatives (except October 2015) with a few false positives until 2021. The probabilities are sufficiently high for the NSR criterion to be reached in March 2014–July 2015 which covers the period of the annexation of Crimea and early troubles in the Donetsk and Luhansk regions.

In Figure 5b, we see a similar picture with a lot of noise in the early period. Although the numbers of applications are low, some changes occur around a number of politically driven events, especially in 2013, with a larger number of crises identified (some correctly, some not). This period is followed by true positives identified for the period December 2013–July 2015. Neither contemporaneous model performs well for the 2021–22 crisis.



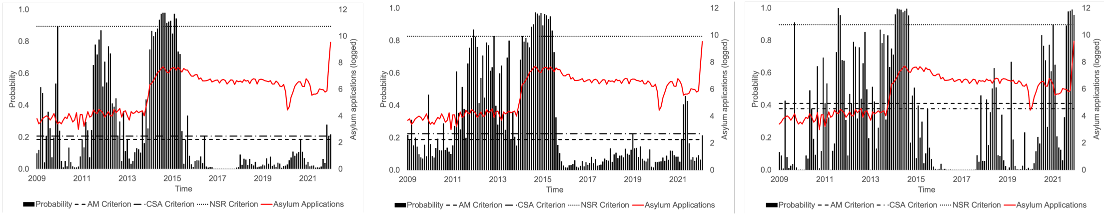
(a) Model T-M1 (b) Model T-M2

Figure 5: EWS results from Ukraine at the current horizon

EWS results for Ukraine for the contemporaneous model. The horizontal lines identify the AM, CSA, and NSR criteria. The vertical bars show the probability at the current period of a crisis occurring with the red line being the natural log of asylum applications. For Model T-M1, the threshold values are: AM=0.2625, CSA=0.2055, NSR=0.7529; Model T-M2: AM=0.3325, CSA=0.3384, NSR=0.9587. Source: Eurostat (asylum data); own elaboration from EWS results.

The six-month horizon models were generally able to predict crisis in February 2022. The results for the first model, defined in (12), are shown in Figure 6a. The 2014 crisis is first identified for November 2013 (false positive), and as with other models, crisis indications continue until mid-2015, with a mix of false positives and negatives in 2013. The crises in February and March 2022 are predicted at a six-month horizon, corresponding to August and September 2021, with a false positive for January 2022. In comparison, the second model, defined in (13) only provides an alert for March 2022, but also for early 2021, as shown in Figure 6b. There are a large number of true positives in 2013, with the crisis beginning in January 2014 correctly identified with a six-month offset, in July 2013, as indicated by a probability surge. The results for the final model, defined in (14), is shown in Figure 6c. With mixed performance in 2013, true positives and true negatives are predicted in 2014 with a series of false positives in 2015. The crisis probability significantly increases in November 2021, with the values from December 2021 onwards exceeding the NSR criterion, which correctly indicates a high likelihood of a crisis even at a six-month horizon.

One important finding is that the models using a six-month horizon were able to predict a crisis correctly, as in this case, it could allow for greater preparedness. With these models, high probabilities for T6-M3 can be seen from November 2021, which could have been seen already in May 2021. The value for 0.998 for January and February 2022, would have seen some signal already in July 2021, which was in advance of the Russian troops being moved towards the Ukrainian border in November 2021, although other political events already started to indicate the potential of a large crisis occurring.



(a) Model T6-M1 (b) Model T6-M2 (c) Model T6-M3

Figure 6: EWS results from Ukraine at a six-month horizon

EWS results for Syria at a six-month horizon. The horizontal lines identify the AM, CSA, and NSR criteria. The vertical bars show the probability at the current period of a crisis occurring with the red line being the natural log of asylum applications. For Model T6-M1, the threshold values are: AM=0.1856, CSA=0.2076, NSR=0.8952; Model T6-M2: AM=0.1889, CSA=0.2247, NSR=0.8272; Model T6-M3: AM=0.4112, CSA=0.3782, and NSR=0.8976. Source: Eurostat (asylum data); own elaboration from EWS results.

4.2.2 Confusion matrix and model performance

The confusion matrix and statistical analysis presented in Table 5 give an intriguing picture emerging from the analysis of all the models. There is no single model that scores highly across all indicators. The rankings for Cohen's Kappa, AUC and MSE do not match as well as they do in the Syrian case study: for Ukraine, the build-up to the crises was quite different. Figure 7 shows the AUC for the models analysed for Ukraine. We can see that as all the curves are above the 45-degree line, there is a greater than 50-50 chance of a crisis being correctly predicted.

In terms of the performance individual models, model T-M1 minimises the type I error (FPR, 1–Specificity), whilst T6-M1 minimises the type II error (FNR, 1–Sensitivity). Model T-M2 maximises Kappa, with T-M1 maximising the AUC, and T-M1 minimising the MSE. The six-month horizon models struggle relatively for the Kappa score. However, for AUC, T6-M3 performs quite well, with T6-M1 close to T-M1 for MSE. The six-month horizon models perform somewhat better based on the AUC and the MSE, but due to the similar magnitude of most measures, all models seem to provide useful information.

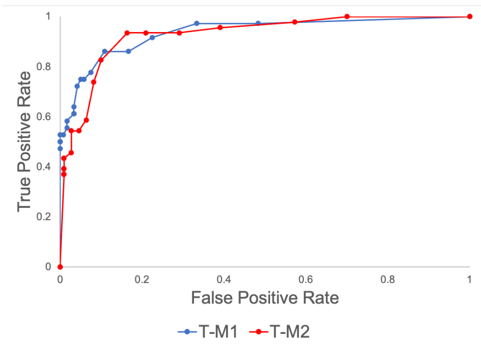
4.2.3 Discussion

One of the problems that comes with identifying crises in Ukraine is the choice of an indicator of a crisis. As seen in the Syrian case, the models with a 12-month growth rate outperform six-month growth rates. However, for all indicators, there is a lot of noise at the beginning of the sample, even though the *number* of asylum applications is low.

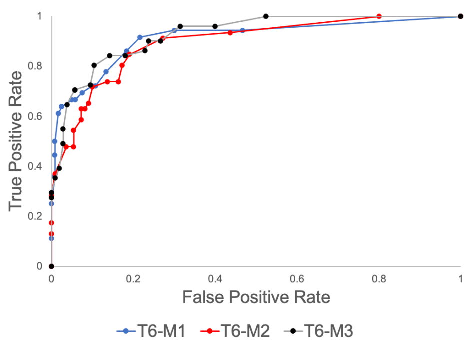
Table 5: Confusion matrix and contingency table proportions

Model	**TP**	**FP**	**FN**	**TN**	**Acc.**	**Prec.**	**Sens.**	**Spec.**	**Kappa**	**AUC**	**MSE**
Signal detection characteristics for the contemporaneous models											
T-M1	31	11	5	11	89.87%	0.7381	0.8611	0.9098	0.7282	0.9367	0.0720
T-M2	42	14	4	98	88.61%	0.7500	0.9130	0.8750	0.7406	0.9276	0.1005
Signal detection characteristics at the six-month horizon											
T6-M1	33	23	3	98	83.43%	0.5893	0.9167	0.8099	0.6080	0.9122	0.0834
T6-M2	38	19	8	93	82.91%	0.6667	0.8261	0.8304	0.6132	0.8950	0.1152
T6-M3	43	13	8	92	86.54%	0.7679	0.8431	0.8762	0.7016	0.9265	0.1022

The confusion matrix details the accuracy of the results by comparing the predicted outcome with the actual outcome. The four outcomes are true positive (TP), false positive (FP), false negative (FN), or true negative (TN). The sum of outcomes differs depending on binary variable selected. Results for the models presented below using the score maximising criterion. The models are evaluated with Accuracy (Acc), Precision (Prec), Sensitivity (Sens), Specificity (Spec), Kappa, Area Under the Curve (AUC) and Mean Squared Error (MSE).



(a) Contemporaneous



(b) 6 Month Horizon

Figure 7: ROC Curve for Ukraine Models

The ROC plots for the eight models analysed in Table 5. The horizontal axis is the False Positive Rate (FPR) and the vertical axis the True Positive Rate (TPR). For each model, the blue line corresponds to the modell in the respective panel of Table 4, red for the model 2, and black for the model 3 (where applicable).

However, tightening these definitions of a crisis then risks losing signal, although probabilities that are close to or exceed the tight NSR criterion tend to pick crisis periods more correctly. This highlights the need for being cautious and using additional contextual information and human input to interpret the data – in this case, related to the changing political climate since 2013 – to deduce that something might be happening.

Does the analysis in Table 5 identify the *best* model? Statistically yes, but from a perspective of predicting a crisis with user input, the results are more nuanced. For instance, some of the most efficient models for Ukraine do not predict the 2022 crisis in February at certain horizons but did so for March 2022. Another consideration is that there is only one full crisis period for the models to learn from, as the second crisis only starts in terms of inflows at the end of February 2022, limiting the models' ability to learn.

5 Conclusion

In this paper, we have proposed a suite of early warning models that managed to successfully trigger some of the expected warnings for the two case studies. The models for these differing crises used varied datasets, including some 'big data' indicators, with modelling techniques adopted largely from finance and macroeconomics to generate warnings with a lead-time of up to six months before the increases in the numbers of applications could be observed. The models produced promising results, and all give a fairly high degree of predictability, far greater than an even chance as shown in the ROC-AUC analysis.

In terms of potential predictors carrying sufficient signal for early warnings, we have looked at an array of various factors and corresponding variables. In our examples, the data came from a range of traditional (economic, geopolitical) and 'big data' sources. The model analysis presented in this paper has demonstrated that no single model can be useful in every context, with different variables being preferable for different applications, situations and countries. While macroeconomic data might not be the first choice for an array of scholars, there are also important insights that can be learned from them. Formal econometric model selection procedures, such as LASSO, can help identify the right set of explanatory variables with sufficient predictive power.

As for individual predictors, exchange rates (in particular, the real effective exchange rates) were found to be quick to react to news. On the news themselves, 'big data' sources such as GDELT provide an excellent, large-scale resource that allows investigating specific actors and circumstances, such as in the case of Ukraine and Russia. Search data, such as Google Trends, introduce another aspect, albeit with limitations due to uneven access to Internet in different locations. Compiling a set of EWS models on a per country basis would provide more accurate results than employing the same model to all situations.

Importantly, a crucial element of building an EWS model involves desk research on the causes of each of the crises, to identify background and context to find why, and how, these conflicts occurred and escalated. Of course, with hindsight, it is always easy to say that there may have been clear signals at the time. Still, while the same or similar geopolitical conditions persist, these signals can help establish warnings for any future crisis, which the results for Syria and Ukraine were able to show to differing extents. In this research, we have tried to strike a balance between human input and data-driven modelling results. Fully grasping the magnitude of the crisis was not something that could be reliably explained by the model alone. In such applications, only human input would be ultimately able to fully confirm the seriousness of the challenge, whilst remaining cognizant of all the ethical and legal aspects involved in relying on models for helping shape the political or humanitarian responses to the crisis of displacement.

References

Avramescu, A., Wiśniowski, A., 2021. Now-casting Romanian migration into the United Kingdom by using Google Search engine data. Demographic Research 45, 1219–1254. doi:10.4054/DemRes.2021.45.40.

Ben-David, A., 2008. About the relationship between ROC curves and Cohen's kappa. Engineering Applications of Artificial Intelligence 21, 874–882.

Bijak, J., Czaika, M., 2020. Black swans and grey rhinos: Migration policy under uncertainty. Migration Policy Practice X, 14–20. doi:10.1186/s40878-022-00284-2.

Bowers, A.J., Zhou, X., 2019. Receiver operating characteristic (ROC) area under the curve (AUC): A diagnostic measure for evaluating the accuracy of predictors of education outcomes. Journal of Education for Students Placed at Risk 24, 20–46.

Böhme, M.H., Gröger, A., Stöhr, T., 2020. Searching for a better life: Predicting international migration with online search keywords. Journal of Development Economics 142, 102347. doi:10.1016/j.jdeveco.2019.04.

Carammia, M., Iacus, S.M., Wilkin, T., 2022. Forecasting asylum-related migration flows with machine learning and data at scale. Scientific Reports 12, 1–16.

Carastathis, A., Spathopoulou, A., Tsilimpounidi, M., 2018. Crisis, What Crisis? Immigrants, Refugees, and Invisible Struggles. Refuge: Canada's Journal on Refugees 34, 29–38.

Crawley, H., 2016. Managing the Unmanageable? Understanding Europe's Response to the Migration 'Crisis'. Human Geography 9, 13–23. doi:10.1177/194277861600900202.

Filippopoulou, C., Galariotis, E., Spyrou, S., 2020. An early warning system for predicting systemic banking crises in the Eurozone: A logit regression approach. Journal of Economic Behavior & Organization 172, 344–363.

Friedman, J., Hastie, T., Tibshirani, R., Narasimhan, B., Tay, K., Simon, N., Qian, J., Yang, J., 2022. Lasso and Elastic-Net Regularized Generalized Linear Models. Technical Report Package 'glmnet". CRAN Repository. Https://cran.r-project.org/web/packages/glmnet/glmnet.pdf. Accessed on 26 October 2022.

Hasse, J.B., Lajaunie, Q., 2021. Package 'EWS'. Technical Report EWS. CRAN Repository. Https://CRAN.R-project.org/package=EWS. Accessed on 22 July 2022.

Jaroszewicz, M., 2019. Years After Crimea's Annexation, Integration of Ukraine's Internally Displaced Population Remains Uneven. Migration Policy Institute. https://www.migrationpolicy.org/article/fyears-after-crimea-annexation-integration-ukraine-internally-displaced-population. Accessed: 28 April 2022.

Juric, T., 2022. Predicting refugee flows from Ukraine with an approach to Big (Crisis) Data: a new opportunity for refugee and humanitarian studies. medRxiv .

Kaminsky, G., Lizondo, S., Reinhart, C.M., 1998. Leading indicators of currency crises. IMF Staff Papers 45, 1–48.

Kauppi, H., Saikkonen, P., 2008. Predicting U.S. Recessions with Dynamic Binary Response Models. The Review of Economics and Statistics 90, 777–791.

Lajaunie, Q., 2021. Nonlinear Impulse Response Function for Dichotomous Models. LEO Working Papers / DR LEO 2852. Orleans Economics Laboratory / Laboratoire d'Economie d'Orleans (LEO), University of Orleans.

Napierała, J., Hilton, J., Forster, J.J., Carammia, M., Bijak, J., 2022. Toward an early warning system for monitoring asylum-related migration flows in Europe. International Migration Review 56, 33–62.

Sahin-Mencutek, Z., Barthoma, S., Gökalp-Aras, N.E., Triandafyllidou, A., 2022. A crisis mode in migration governance: comparative and analytical insights. Comparative Migration Studies 10.

Wu, C., Gerber, M.S., 2018. Forecasting civil unrest using social media and protest participation theory. IEEE Transactions on Computational Social Systems 5, 82–94.

Zelinska, O., 2017. Ukrainian Euromaidan protest: Dynamics, causes, and aftermath. Sociology Compass 11, e12502.