Sovereign Defaults: a Panel Binary Conditional Forecasting Approach

Ana Beatriz Galvão, Michael McCracken, Michael Owyang

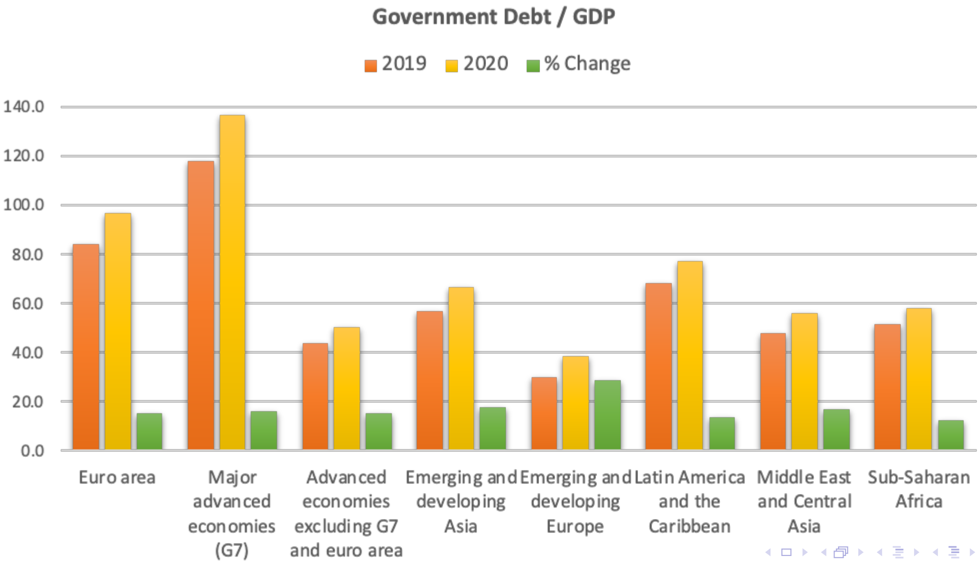
Warwick, CEPR/St Louis Fed

The views expressed here are the authors'alone do not reflect those of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

February 2022

Covid-19 Pandemic and Sovereign Debt

- Government Debt to GDP ratios increased substantially in 2020.



Covid-19 Pandemic and Sovereign Defaults

- Credit rating agencies (such as Moody's) described 2020 as a record year for the number of countries in default: Argentina, Belize, Ecuador, Lebanon, Suriname, Zambia (in addition to Venezuela).

Covid-19 Pandemic and Sovereign Defaults

- Credit rating agencies (such as Moody's) described 2020 as a record year for the number of countries in default: Argentina, Belize, Ecuador, Lebanon, Suriname, Zambia (in addition to Venezuela).
- What are the quantitative effects of the Covid-19 pandemic on the probability of default in the years to come?

Our Approach: The P-Qual-VAR

- A panel VAR approach where intercepts are country-specific.

Our Approach: The P-Qual-VAR

- A panel VAR approach where intercepts are country-specific.
- The underlined latent risk of the event is a VAR endogenous variable, and it is directly linked to the binary variable representing the event history.

Our Approach: The P-Qual-VAR

- A panel VAR approach where intercepts are country-specific.
- The underlined latent risk of the event is a VAR endogenous variable, and it is directly linked to the binary variable representing the event history.
- The model allows for the computation of dynamic effects of shocks to the event probabilities.

Our Approach: The P-Qual-VAR

- A panel VAR approach where intercepts are country-specific.
- The underlined latent risk of the event is a VAR endogenous variable, and it is directly linked to the binary variable representing the event history.
- The model allows for the computation of dynamic effects of shocks to the event probabilities.
- Conditional forecasts are applied to evaluate the pandemic effects on the predicted probability of default.

Previously Considered Alternative Approaches: Historical Analysis

- Historical analyses of how default events are linked to sovereign debt and other economic factors (Reinhart and Rogoff, 2011; Borensztein and Panizza, 2009; Esteves et al, 2021).

Previously Considered Alternative Approaches: Historical Analysis

- Historical analyses of how default events are linked to sovereign debt and other economic factors (Reinhart and Rogoff, 2011; Borensztein and Panizza, 2009; Esteves et al, 2021).
- Quantitative models of sovereign defaults and the economy (surveys: Tomz and Wright, 2013 and Yue and Bei, 2018).

Previously Considered Alternative Approaches: Historical Analysis

- Historical analyses of how default events are linked to sovereign debt and other economic factors (Reinhart and Rogoff, 2011; Borensztein and Panizza, 2009; Esteves et al, 2021).
- Quantitative models of sovereign defaults and the economy (surveys: Tomz and Wright, 2013 and Yue and Bei, 2018).
- Tree-based models to predict whether a country is in default in a given year based on a large set of economic and political variables (Manasse and Roubini, 2009; and Savona and Vezzoli, 2015).

Previously Considered Alternative Approaches: Historical Analysis

- Historical analyses of how default events are linked to sovereign debt and other economic factors (Reinhart and Rogoff, 2011; Borensztein and Panizza, 2009; Esteves et al, 2021).
- Quantitative models of sovereign defaults and the economy (surveys: Tomz and Wright, 2013 and Yue and Bei, 2018).
- Tree-based models to predict whether a country is in default in a given year based on a large set of economic and political variables (Manasse and Roubini, 2009; and Savona and Vezzoli, 2015).
- Stylised facts: The unconditional probability of default is 2% (1820-2012) but has increased to 4% recently (1980-2012). The average default duration is 8 years, so the fraction of time spent in default is 18%.

Previously Considered Alternative Approaches: Historical Analysis

- Historical analyses of how default events are linked to sovereign debt and other economic factors (Reinhart and Rogoff, 2011; Borensztein and Panizza, 2009; Esteves et al, 2021).
- Quantitative models of sovereign defaults and the economy (surveys: Tomz and Wright, 2013 and Yue and Bei, 2018).
- Tree-based models to predict whether a country is in default in a given year based on a large set of economic and political variables (Manasse and Roubini, 2009; and Savona and Vezzoli, 2015).
- Stylised facts: The unconditional probability of default is 2% (1820-2012) but has increased to 4% recently (1980-2012). The average default duration is 8 years, so the fraction of time spent in default is 18%.
- The cost of sovereign defaults on GDP is 1.6% one year after the event (Esteves et al, 2021).

Previously Considered Alternative Approaches: Historical Analysis

- Historical analyses of how default events are linked to sovereign debt and other economic factors (Reinhart and Rogoff, 2011; Borensztein and Panizza, 2009; Esteves et al, 2021).
- Quantitative models of sovereign defaults and the economy (surveys: Tomz and Wright, 2013 and Yue and Bei, 2018).
- Tree-based models to predict whether a country is in default in a given year based on a large set of economic and political variables (Manasse and Roubini, 2009; and Savona and Vezzoli, 2015).
- Stylised facts: The unconditional probability of default is 2% (1820-2012) but has increased to 4% recently (1980-2012). The average default duration is 8 years, so the fraction of time spent in default is 18%.
- The cost of sovereign defaults on GDP is 1.6% one year after the event (Esteves et al, 2021).
- Limited knowledge of dynamic effects: use of annual data.

Previously Considered Alternative Approaches: CDS Spreads

- Global financial factors have a critical role in driving fluctuations in CDS sovereign spreads (Longstaff et al, 2011). Default probabilities can be computed based on CDS spreads (Lucas et al, 2014). They have been used to evaluate the impact of the Covid-19 on the sovereign risk of European countries (Augustin et al, 2021).

Previously Considered Alternative Approaches: CDS Spreads

- Global financial factors have a critical role in driving fluctuations in CDS sovereign spreads (Longstaff et al, 2011). Default probabilities can be computed based on CDS spreads (Lucas et al, 2014). They have been used to evaluate the impact of the Covid-19 on the sovereign risk of European countries (Augustin et al, 2021).
- A disadvantage of using CDS sovereign spreads to understand sovereign default dynamics is they are only available for advanced and emerging economies from mid-1990's. The list of countries considered by Augustin (2018), for example, does not include Belize, Ecuador, Suriname and Zambia, which are countries that have partially defaulted in their debt in 2020.

The P-Qual-VAR applied to Sovereign Default I

- Default Probabilities are estimated at quarterly frequency using the sovereign default events from Asonuma and Trebesch (2016). The event end date depends of when the default was resolved. Unconditional probability of 18%.

The P-Qual-VAR applied to Sovereign Default I

- Default Probabilities are estimated at quarterly frequency using the sovereign default events from Asonuma and Trebesch (2016). The event end date depends of when the default was resolved. Unconditional probability of 18%.
- Global Variables: global industrial activity and commodity market effects, we use Killian's (2009, 2019) index; the one-year US Treasury bill rate to measure possible links between advanced-economy monetary policy and sovereign defaults in emerging economies; US excess bond premium to measure global financial cycles (Gilchrist et al., 2021).

The P-Qual-VAR applied to Sovereign Default I

- Default Probabilities are estimated at quarterly frequency using the sovereign default events from Asonuma and Trebesch (2016). The event end date depends of when the default was resolved. Unconditional probability of 18%.
- Global Variables: global industrial activity and commodity market effects, we use Killian's (2009, 2019) index; the one-year US Treasury bill rate to measure possible links between advanced-economy monetary policy and sovereign defaults in emerging economies; US excess bond premium to measure global financial cycles (Gilchrist et al., 2021).
- Domestic Variables: year-on-year GDP growth, government debt and sovereign debt ratios to GDP, and the year-on-year growth in the real effective exchange rate.

The P-Qual-VAR applied to Sovereign Default II

- Measurement of Dynamic Effects: (i) from global and domestic variables to the probability of default, and (ii) from exogenous changes in the probability of default to domestic variables.

The P-Qual-VAR applied to Sovereign Default II

- Measurement of Dynamic Effects: (i) from global and domestic variables to the probability of default, and (ii) from exogenous changes in the probability of default to domestic variables.
- Computation of conditional forecasts which are useful to evaluate how the 2020 pandemic affected the future probability of default .

P-Qual-VAR: Binary Variable

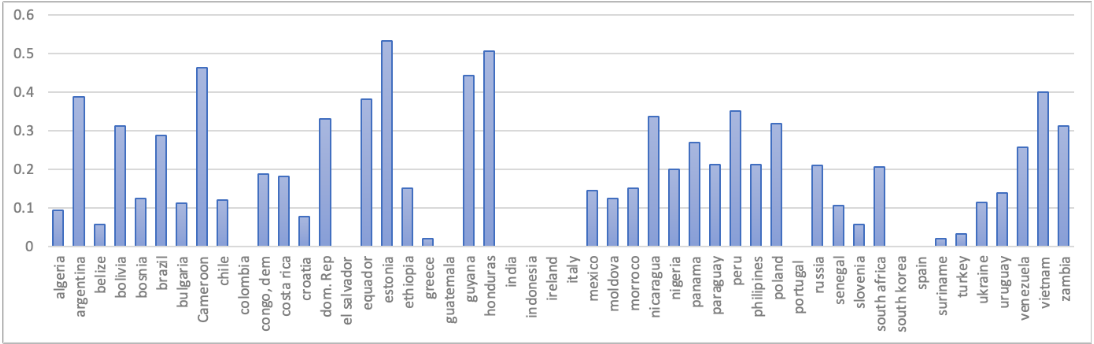
- S_{ct} is a binary indicator variable that is equal to 1 if the event of interest is observed in the country c at time t, and zero otherwise. $t = 1, ..., T$ and $c = 1, ..., N$
- The latent variable z^*_{ct} is defined such that:

$$\begin{aligned} \textit{if } S_{ct} &= 1 \text{ then } z^*_{ct} \geq 0 \\ \textit{if } S_{ct} &= 0 \text{ then } z^*_{ct} < 0. \end{aligned}$$

- z^*_{ct} is a continuous counterpart of the binary variable, measuring the risk of occurrence of the event.

Unconditional Default Probabilities

- $N = 50$: 22 Central + South America, 5 Asia, 8 Africa/Middle East and 15 European countries (EE and SE).
- $S_{ct} = 1$ if country c has default in part of their sovereign debt (or started a preemptive restructuring) at quarter t, or has still not yet resolved an earlier default event, as compiled by Asonuma and Trebesch (2016, updated).



Notes: The sample period is 1980Q1-2020Q4 with the exception of seven Eastern European countries with sample period 1993Q4-2020Q4.

P-Qual-VAR: Other Endogenous Variables

- X_{ct} is the $m \times 1$ vector of domestic variables for country c.
- Based on predictive literature and data availability, $m = 4$:
  1. year-on-year GDP growth (WEO, Eurostat, other sources);
  2. government debt to GDP ratio (IMF Global Debt Database);
  3. external debt to GDP ratio (IMF/IFS)
  4. REER (IMF/IFS + Darvas database).
- W_t is the $w \times 1$ vector of global variables. $w = 3$.
  1. Excess Bond Premium (EBP) as measure of global financial risk (Gilchrist et al, 2021);
  2. Index of Global Economic Activity (GEA) (Killian, Dallas Fed) to measure global industrial activity;
  3. US 1-year Treasury Bill interest rates to measure advanced economies monetary policy.

P-Qual-VAR: Specification

- Assume $y_{ct} = [W'_t, X'_{ct}, z^*_{ct}]'$, so the number of endogenous variables is $k = 1 + m + w$.
- The P-Qual-VAR is:

$$y_{ct} = \mathbf{b}_c + \mathbf{B}_1 y_{ct-1} + ... + \mathbf{B}_p y_{ct-p} + \mathbf{u}_{ct}; \mathbf{u}_{ct} \sim N(0, \mathbf{\Sigma}),$$

  where $\mathbf{b}_c$ are country-specific intercepts (individual effects).
- The probability of default is $\Pr ob(z^*_{ct} \geq 0) = \Phi(z^*_{ct} / \sigma_{z^*_c})$, where Φ is the standard normal CDF.
- The variance-covariance matrix includes the restriction that $var(u^z_{ct}) = 1$, as required for identification.

P-Qual-VAR: estimation I

- The sampler has three blocks:

1. the VAR parameters: $\mathbf{B}, \mathbf{\Sigma} | \mathbf{b}_c, \{z^*_{ct}\}_{t=1;c=1}^{T;N}$;
2. the country-specific intercepts: $\mathbf{b}_c | \mathbf{B}, \mathbf{\Sigma}, \{z^*_{ct}\}_{t=1,c=1}^{T,N}$;
3. the latent variable: $z^*_{ct} | \mathbf{b}_c, \mathbf{B}, \mathbf{\Sigma}$ for $t = 1, ..., T$ and $c = 1, ..., N$.

P-Qual-VAR: estimation II

- For the first block, we use the following representation:

$$y_{ct} - \mathbf{b}_c = \mathbf{B}_1 y_{ct-1} + ... + \mathbf{B}_p y_{ct-p} + \mathbf{u}_{ct},$$

which can be written as:

$$\tilde{y}_{ct} = \Psi_{ct} \mathbf{B} + \mathbf{u}_{ct},$$

where $\tilde{y}_{ct} = (y_{ct} - \mathbf{b}_c)'$ is $1 \times k$ vector, $\Psi_{ct} = [y'_{ct-1}, ..., y'_{ct-p}]$ is a $1 \times kp$ vector and $\mathbf{B} = (\mathbf{B}_1, ..., \mathbf{B}_p)'$ is a $kp \times k$ matrix. The system of equations can be stacked over all countries and time periods as:

$$\underset{TN \times k}{\tilde{y}} = \underset{TN \times kp}{\Psi} \underset{kp \times k}{\mathbf{B}} + \underset{TN \times k}{\mathbf{U}}.$$

P-Qual-VAR: estimation III

- Then assuming Minnesota Normal/Wishart priors for $\mathbf{B}$ and $\boldsymbol{\Sigma}$, we can follow Banbura et al (2010) and implement the priors by adding dummy observations.
- Conditional draws for $\mathbf{B}$ and $\boldsymbol{\Sigma}$ are then obtained using the Normal and Wishart distributions with parameters obtained with the close form solution described in Banbura et al (2010). We fix overall prior tightness to 0.02.

P-Qual-VAR: estimation IV

- The priors for country-specific intercepts are common across countries: $\mathbf{b}_c|V \sim N(0, \mathbf{V})$ and $\mathbf{V}^{-1} \sim W(v_0, \mathbf{V}_0)$.
- Then the conditional posteriors for each country c is $\mathbf{b}_c \sim N(\overline{\mathbf{b}}_c, \overline{\mathbf{V}}_c)$, where

$$
\begin{aligned}
\overline{\mathbf{V}} &= [\mathbf{V}_0^{-1} + \sum_{c=1}^{N} \mathbf{b}_c \mathbf{b}_c']^{-1} \\
\overline{\mathbf{V}}_c &= [T\Sigma^{-1} + \overline{\mathbf{V}}^{-1}]^{-1} \\
\overline{\mathbf{b}}_c &= \overline{\mathbf{V}}_c [\Sigma^{-1} \zeta_c i_T],
\end{aligned}
$$

where $\zeta_{ct} = y_{ct} - \mathbf{B}_1 y_{ct-1} - ... - \mathbf{B}_p y_{ct-p}$ and i_T is a $T \times 1$ vector of 1s, ζ_c is $k \times T$ by getting all ζ_{ct} for country c.

P-Qual-VAR: estimation V

- Conditional on $\mathbf{b}_0$,$\mathbf{B}_1, ..., \mathbf{B}_p, \mathbf{\Sigma}$ and the observed data (W_t, X_{ct}, S_{ct}), we use the method in Dueker (2005) and extended by McCracken et al (2021) to obtain draws for z^*_{ct} for $t = 1, ..., T$ for each country c.
- [The filtering uses a Metropolis strategy for $t = 1, ..., p$, for $t = p + 1, ..., T - p$ the Dueker (2005) closed-form conditional densities, and for $t = T - p + 1, ..., T$ the closed-form densities for conditional forecasts].
- The draws for z^*_{ct} $t = 1, ..., T$ are obtained using each country VAR.
- Then the P-Qual-VAR can be easily estimated over unbalanced panels (and we have 7 countries with data availability from 1993Q4).

P-Qual-VAR: Dynamic Responses of Event Probabilities

- We compute dynamic responses for horizons $r = 0, 1, ..., R$ as:

$$Prob[S_{t+r} = 1|\Omega_t, \varepsilon_j] - Prob[S_{t+r}|\Omega_t] =$$
$$Prob[S_{t+r} = 1|\Omega_t, \varepsilon_j = \mathbf{b}_0] - Prob[S_{t+r} = 1|\Omega_t, \varepsilon_j = \mathbf{0}] =$$
$$\Phi(\hat{z}^*_{t+r}/\hat{\sigma}^*_{z,r}) - 0.5,$$

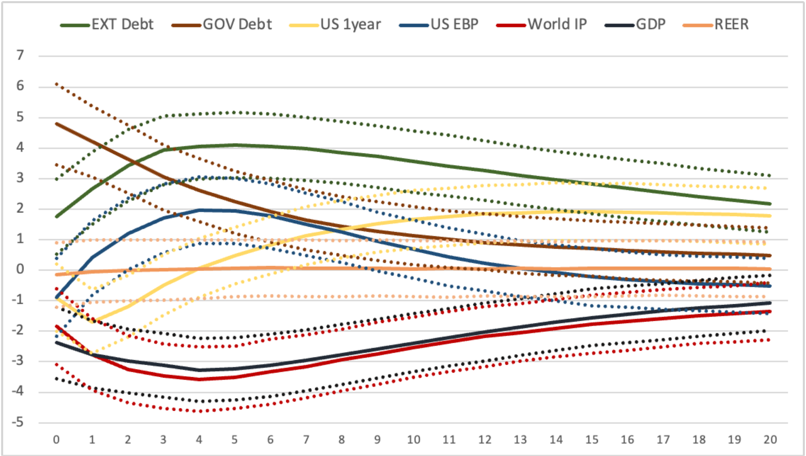
- where $\hat{z}^*_{t+r}$ and $\hat{\sigma}^*_{z,r}$ are computed using simulated paths conditional on $\mathbf{b}_0$ using VAR:

$$y_t = \mathbf{B}_1 y_{t-1} + ... + \mathbf{B}_p y_{t-p} + \mathbf{u}_t; \mathbf{u}_t \sim N(0, \Sigma).$$

- The impact of the shock j on the endogenous variables may be computed by applying a Cholesky decomposition to the variance-covariance matrix.
- The responses is computed for each kept conditional posterior draw of the parameters. We present the mean response and 68% credible intervals.

P-Qual-VAR: Effects of Global and Domestic variables on Default Prob.

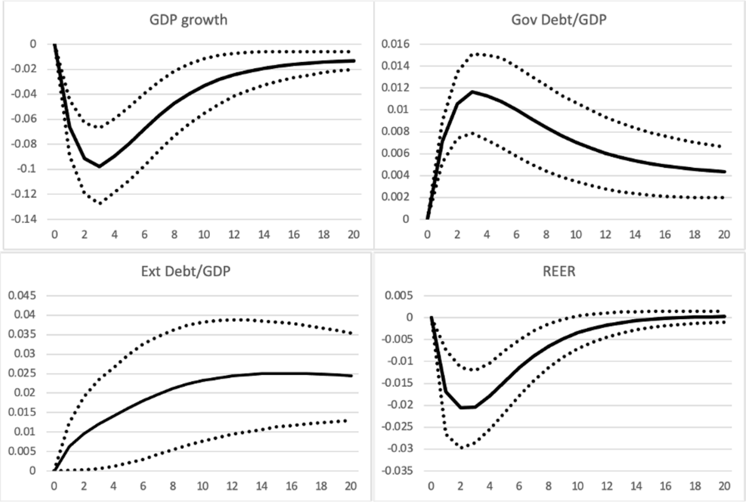
- Using full sample. Values in p.p..



- Dotted lines are 68% credible intervals. Recursive Cholesky-based impact effects.

P-Qual-VAR: Effects of an increase in Default Prob.

- Response of Domestic Variables to a 33.3 p.p. increase in Default Probability.



- These are computed assuming that default probabilities have zero impact effects on global and domestic variables. Dotted lines are 68% credible intervals.

P-Qual-VAR:Forecasting I

- Using the country-specific Qual-VAR models conditional on $\mathbf{b}_c, \mathbf{B}, \boldsymbol{\Sigma}$ and $\{z^*_{ct}\}^{\tau}_{\tau-p}$, we can compute multi-step forecasts $y_{c\tau+h}$ for $h = 1, ..., H$ for all endogenous variables $(W_{\tau+h}, X_{c\tau+h}, z^*_{c\tau+h})$.
- Forecasts for the event probability are then $P_{c\tau+h} = Prob(S_{c\tau+h} = 1) = \Phi(z^*_{c\tau+h}/\sigma_{z^*_{ch}})$.
- A caveat is that predictions for $W_{\tau+1}, ..., W_{\tau+h}$ will differ for each c.

P-Qual-VAR:Forecasting II

- We compute forecasts from the P-Qual-VAR in two steps:
  1. Compute the country-specific multi-step unconditional forecasts for the global variables.
  2. Compute forecasts for for $X_{c\tau+h}$, $z^*_{c\tau+h}$ and $P_{c\tau+h}$ conditional on $\overline{W}_{\tau+1}, ..., \overline{W}_{\tau+h}$, which are averages of country-specific forecasts.
- Conditional forecasts use the algorithm in McCracken et al (2021), which is based on Antolin-Diaz et al (2020).
- The forecasting algorithm is computed at each kept posterior draw of the parameters. The predicted probabilities are the mean across these draws (5,000).

Event Forecasting: Measures of Accuracy I

- The logarithm score (LS) is computed for each country over M observations as:

$$LS(c,h) = \frac{1}{M}\sum_{\tau=1}^{M} \ln\left|1 - S_{c\tau+h} - P_{c\tau+h}\right|.$$

- For the panel, the logscore is:

$$LS(h) = \frac{1}{N}\sum_{c=1}^{N} LS(c,h).$$

Event Forecasting: Measures of Accuracy II

- Bouallegue et al. (2019) recommends the diagonal of the elementary score (DES) when binary events are rare but have high-impact consequences and false positives do not cause large losses. In contrast to the AUROC, the DES is a proper score:

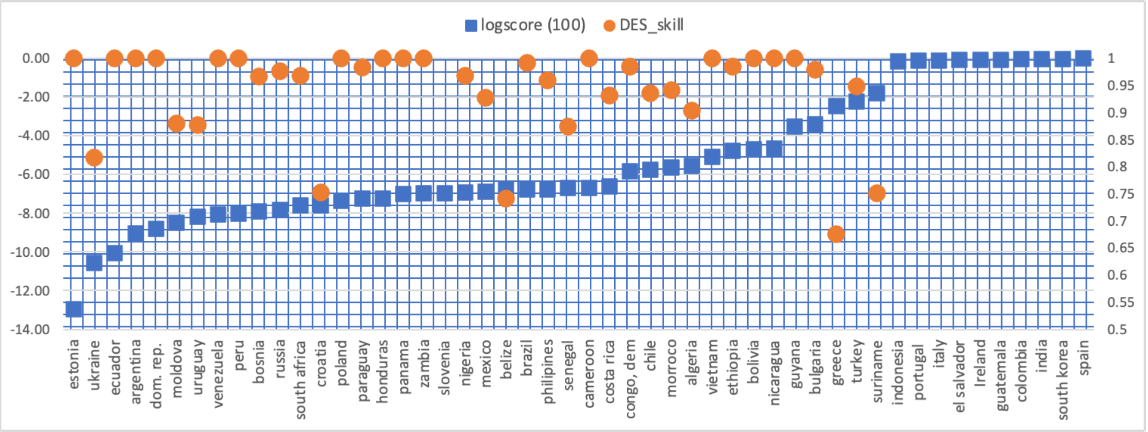
$$DES(c,h) = \frac{1}{M}\sum_{\tau=1}^{M} \pi_c \mathbf{I}[P_{c\tau+h} > \pi_c](1 - S_{c\tau+h}) + (1 - \pi_c)\mathbf{I}[P_{c\tau+h} \leq \pi_c] S_{c\tau+h},$$

where $\mathbf{I}[.]$ is an indicator function, and π_c is the unconditional probability of the event ($\pi_c = \frac{1}{M}\sum_{\tau=1}^{M} S_{c\tau+h}$).

Event Forecasting: Measures of Accuracy III

- If $\pi_c < 0.5$, false positives are given more weight than false negatives. If all events are classified, $DES = 0$. If all are missed, $DES = \pi_c(1 - \pi_c)$.
- For the unconditional probability forecast, $DES = \pi_c(1 - \pi_c)$. So to measure gains in comparison to the unconditional forecast, we use $DESS(c, h) = 1 - \frac{DES(c,h)}{\pi_c(1-\pi_c)}$. $DESS = 1$ for all events classified, and $DESS = 0$ for no correct classification.

Accuracy of In-Sample Default Probabilities



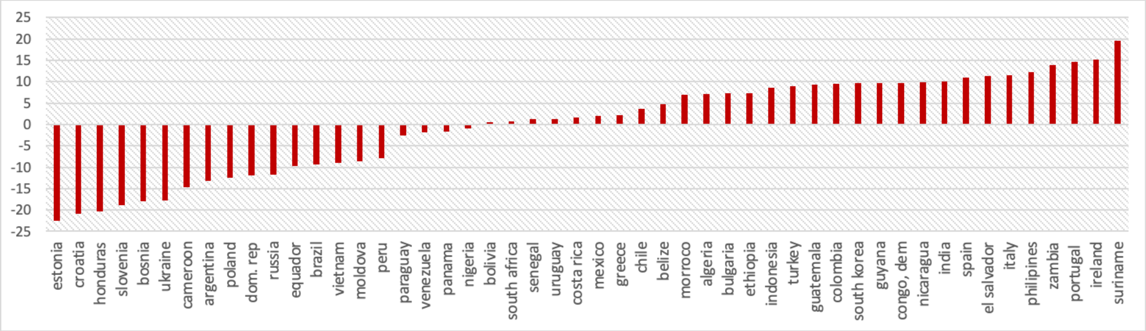
Logscore values were multiplied by 100 and shown on the left axis. The DESS are in the right axis. The sample period is 1980Q1-2020Q4 with the exception of seven Eastern European countries with sample period 1993Q4-2020Q4. $M = T - p$

P-Qual-VAR: OOS

- Competitor: a pooled Dynamic Probit with one lag for all predictors (P-Qual-VAR has $p = 4$) estimated with Bayesian methods by data augmentation. Only $h = 1$.
- Forecast origins: 2010Q1-2019Q4 ($M = 40$) for $h = 1, ..., 4$.
- We computed the DESS by pooling probability forecasts and default events across all countries.
- For $h = 1$, the DESS is 0.28 for P-Qual-VAR but 0 for DynProbit. For $h = 4$, the P-Qual-VAR DESS declines to 0.06.
- If the competitor is the country-specific in-sample (up to 2009Q4) unconditional probability, we can show that the P-Qual-VAR is significantly more accurate using a pooled t-test.

OOS:P-Qual-VAR vs DynProbit, $h = 1$

- Logscore % Gains (P-Qual-VAR to DynProbit)
- Origins: 2010Q1-2019Q4;



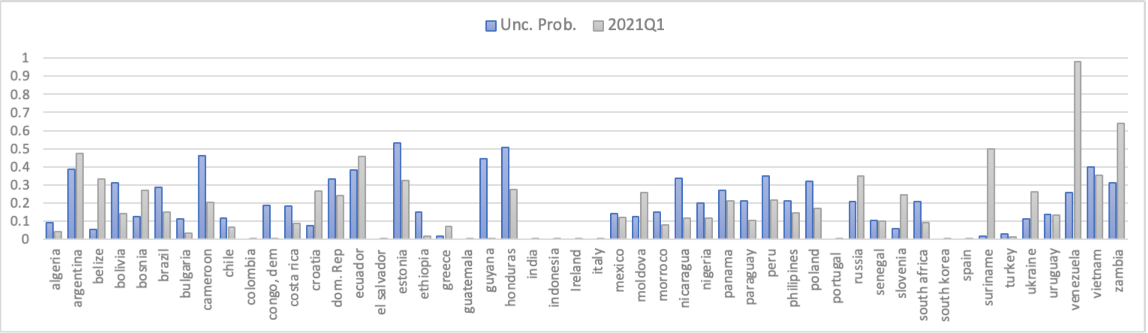
- The average gain is 2.9% if EE countries removed. This is a statistically significant logscore gain using pooled t-test.

Conditional Forecasts to Evaluate Pandemic Effects on Default Probabilities

- We consider differences between "scenario-based" and "baseline" predictive probabilities for 2021Q4 and 2022Q4.
- In the baseline, forecasts for 2022Q4 is a $h = 8$ forecast.
- But for the scenario-based, 2022Q4 is a $h = 12$ forecast conditional on different assumptions on data for 2020Q1-2020Q4.

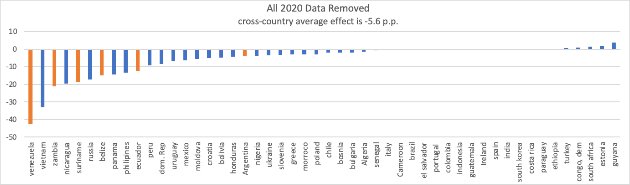
1. All 2020 endogenous variables values are excluded; Differences are useful to evaluate the impact of the changes in the 2020 pandemic year.
2. Only global variables values for 2020 are included; Differences are useful to evaluate the contribution of domestic variables.
3. Global and domestic variables values for 2020 are included; Differences are useful to evaluate exogenous changes in default events - policy or luck.

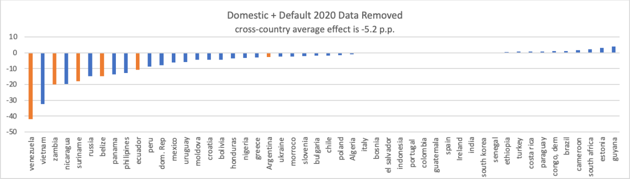
Baseline 2021Q1 Default Probability Forecasts I

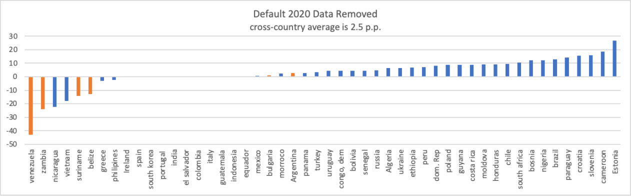


- The baseline probability forecasts using all data up to 2020. The predicted probabilities are the average across 5,000 predicted probabilities computed for each set of posterior draws.

Pandemic Effects in the 2021Q4 Forecasted Probabilities

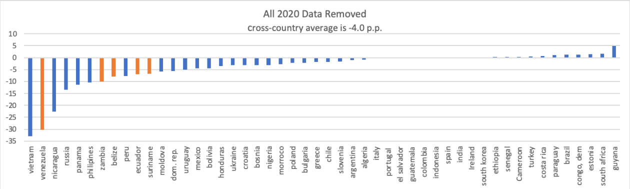


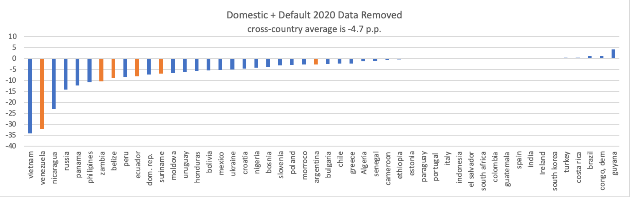


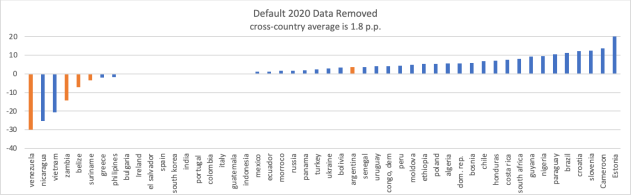


- Negative effects: the 2020 data positively affected default probs in 2021Q4. Countries in

Pandemic Effects in the 2022Q4 Forecasted Probabilities







- Negative effects: the 2020 data positively affected default probs in 2022Q4. Countries in

What did we learn?

- Default probabilities are predicted by both global and domestic factors. In the short run, government debt, and global and domestic economic activity have a crucial role. In the long run, the size of the sovereign debt and US interest rates are more important.
- Exogenous changes in default probabilities have adverse short-term effects on domestic growth, lead to short-lived domestic currency depreciation, and increase external debt in the long term.
- The Covid-19 pandemic marginal effects on the probability of default in 2021-2022 vary across countries. Impacts above 10 p.p. were estimated for 10 countries.

Conclusions

- The P-Qual-VAR was proposed to estimate dynamic effects and conditional forecasts for binary time series variables using a panel of countries.

Conclusions

- The P-Qual-VAR was proposed to estimate dynamic effects and conditional forecasts for binary time series variables using a panel of countries.
- The P-Qual-VAR was applied to sovereign default probabilities using quarterly data from 1980 to 2020 for 50 countries.

Conclusions

- The P-Qual-VAR was proposed to estimate dynamic effects and conditional forecasts for binary time series variables using a panel of countries.
- The P-Qual-VAR was applied to sovereign default probabilities using quarterly data from 1980 to 2020 for 50 countries.
- The model provides evidence of the relative importance of global and domestic factors in predicting default probabilities.

Conclusions

- The P-Qual-VAR was proposed to estimate dynamic effects and conditional forecasts for binary time series variables using a panel of countries.
- The P-Qual-VAR was applied to sovereign default probabilities using quarterly data from 1980 to 2020 for 50 countries.
- The model provides evidence of the relative importance of global and domestic factors in predicting default probabilities.
- An assessment of the impact of the pandemic suggests that good policies/good luck in 2020 have a role in reducing default probabilities for 2021 and 2022 for some countries.