

Indicadores avançados de crises bancárias: exploração de novos dados e instrumentos¹

António Antunes² | Diana Bonfim³ | Nuno Monteiro⁴ | Paulo M. M. Rodrigues⁵

RESUMO

A previsão de eventos raros constitui um desafio, sobretudo se estes forem determinados por fatores muito distintos e assumirem diferentes características. Neste artigo explora-se a dimensão dinâmica de modelos de escolha discreta para melhorar a qualidade das previ-

sões subjacentes a modelos avançados para crises bancárias sistémicas. Os resultados mostram que a inclusão desta componente dinâmica nos modelos melhora de forma significativa a qualidade dos resultados.

1. Introdução

Será possível prever a próxima crise bancária com exatidão? Muito provavelmente não. Por um lado, pode argumentar-se que é econometricamente complexo prever eventos com uma natureza tão rara, muitas vezes determinados por diferentes causas e com diferentes consequências. Por outro lado, se fosse possível prever com rigor a emergência de uma crise bancária com uma antecipação razoável, os decisores de política tomariam todas as medidas necessárias para evitar a materialização de tal crise, o que implicaria uma falha da metodologia de previsão.

O objetivo deste artigo será contribuir para melhorar o conjunto de indicadores avançados disponíveis para os decisores de política. Ao longo das últimas décadas, surgiram inúmeras e diversificadas contribuições para esta literatura, com o intuito de ajudar a compreender os principais determinantes de crises financeiras, bem como auxiliar as autoridades a prever a próxima crise. Uma parte significativa desta literatura incide sobre crises cambiais, sobretudo em economias de mercado emergentes (Krugman, 1979, Obstfeld, 1986, Burnside *et al.*, 2004, Chang e Velasco, 2001). Em todo o caso, é muito frequente que as crises cambiais sejam acompanhadas por crises bancárias, conforme discutido por Kaminsky e Reinhart (1999). Quando uma crise financeira é caracterizada por perturbações e perdas significativas no sistema bancário, os efeitos negativos sobre a economia são habitualmente mais pronunciados e duradouros (Cecchetti *et al.*, 2009, Jordà *et al.*, 2010, 2012).

Apesar de cada crise parecer ser diferente e única (Reinhart e Rogoff, 2011), este artigo explora os elementos comuns numa base de dados de crises bancárias sistémicas na Europa. O seu principal contributo consiste na exploração da componente dinâmica de séries temporais das variáveis dependentes e independentes. A análise efetuada permite concluir que a utilização de um *probit* dinâmico contribui para melhorar a qualidade de previsão de modelos avançados, tanto dentro como fora da amostra (*in-* e *out-of-sample*)

Este artigo está organizado da seguinte forma. Na secção 2 descrevem-se os dados e apresentam-se os modelos e a metodologia de estimação utilizada. Na secção 3 discutem-se os principais resultados, analisa-se a capacidade de previsão dos modelos e apresentam-se alguns testes de robustez. Por fim, na secção 4 sumariam-se os principais resultados.

2. Dados e metodologia

2.1. Dados

Este artigo foi desenvolvido no âmbito do *workshop* do BCE sobre *Early Warning Tools for Supporting Macroprudential Policies*. Foi definido um conjunto de regras e construída uma base de dados comum como ponto de partida. Aos participantes era permitida a manipulação desta base de dados de forma a criar novas variáveis, bem como a utilização de outras fontes públicas de informação. A única restrição em termos de dados utilizados era que os cálculos não utilizassem informação indisponível no momento a que se refere cada observação para o cálculo de tendências e filtros, isto é, simulando uma evolução dos dados em tempo real e sem qualquer interferência de informação futura existente na base de dados. Desta forma seria possível replicar, de forma mais realista, a informação relevante para a tomada de decisões de política disponível a cada momento do tempo. A utilização de defasamentos temporais de variáveis ou o cálculo de tendências de forma recursiva são exemplos de alterações à estrutura contemporânea dos dados originais que seriam admissíveis⁶.

Apesar dos participantes poderem usar diferentes metodologias e fontes de dados, a série de crises bancárias era comum, de forma a garantir a comparabilidade dos resultados. Como evidenciado por Chaudron e de Haan (2014), existem diferenças assinaláveis entre diferentes bases de dados de crises bancárias que se encontram disponíveis publicamente. Boyd *et al.* (2009) defendem que, muitas vezes, as datas das crises refletem intervenções governamentais para gerir dificuldades no sector bancário e não o surgimento destas dificuldades. De forma a garantir a melhor qualidade possível para esta variável crítica, usou-se a base de dados de crises bancárias sistémicas elaborada pelo banco central da República Checa (Babecky *et al.*, 2012). A base de dados de crises conta também com contribuições do grupo *Heads of Research of the Eurosystem* e foi recentemente atualizada com contribuições do *ESRB/IWG Expert Group* (para mais detalhes, ver Detken *et al.*, 2014). Esta base de dados usa duas definições de crise: uma considerando apenas crises bancárias que efetivamente ocorreram e outra que também inclui episódios de elevada vulnerabilidade nos quais se considera, *ex-post*, ter havido motivos para implementar instrumentos macroprudenciais, ainda que não tenha sido registada nenhuma crise⁷. Embora para o exercício apresentado no *Workshop on Early Warning Tools and Tools for Supporting Macroprudential Policies* do BCE se tenha utilizado esta última definição de crise, neste artigo considerar-se-á a definição mais abrangente.

No *workshop* referido anteriormente realizou-se uma *horse race* entre várias metodologias, como discutido em Alessi *et al.* (2014). De forma a permitir comparabilidade entre os resultados das diferentes técnicas, todos os participantes apresentaram os valores das matrizes de contingência e da área sob a curva ROC (AUROC) dos seus modelos⁸.

O exercício foi conduzido em três horizontes diferentes: período completo, período distante e período próximo. O período completo incorpora os últimos 4 a 20 trimestres antes da crise, podendo ser dividido em dois subperíodos: o período distante (12 a 20 trimestres anteriores), que será mais indicado para funcionar como indicador avançado e dar mais tempo de reação para a implementação de medidas; e o período próximo (4 a 12 trimestres anteriores), que será o período imediatamente anterior ao evento de crise, e durante o qual os sinais poderão ser mais fortes e as medidas poderão ter de ser tomadas mais prontamente. Oet *et al.* (2010) apresentam uma fundamentação semelhante para esta estrutura.

Por fim, os participantes deveriam complementar a apresentação dos seus modelos com um exercício *out-of-sample*. Dois exercícios foram sugeridos: um em que seria excluída a crise financeira global e outro em que seriam excluídos da amostra a Dinamarca, a Finlândia e a Suécia, onde no final dos anos 1980/início dos anos 1990 ocorreu uma crise bancária.

Para além de dados sobre crises bancárias, a base de dados incluía diversas variáveis macroeconómicas e financeiras provenientes de várias fontes: crédito concedido ao sector privado (BIS e FMI); preços de imóveis (UE, BIS e OCDE); preços de ativos (UE e FMI); PIB real e nominal (UE, FMI); rácios de serviço da dívida (BIS, BCE e EUROSTAT; cálculos do BCE baseados na metodologia de Drehmann e Juselius, 2012). Adicionalmente, algumas variáveis do sistema bancário também se encontram na base de dados: margem financeira (OCDE), resultados líquidos antes de impostos (OCDE), capital e reservas (OCDE, UE), rácio de alavancagem (UE), e ativos totais (UE). Tentou-se maximizar a informação disponível, pelo que algumas séries de determinadas fontes foram combinadas com dados de outras fontes disponíveis na base de dados. A maioria das variáveis na amostra apresenta séries com valores trimestrais, tendo a série mais longa informação entre 1970T1 e 2010T4. Os resultados deste artigo foram obtidos utilizando uma versão atualizada (com recurso à *Thomson Reuters*) da base de dados originalmente fornecida pelo BCE. Quando as fontes de dados não eram as mesmas, as séries foram prolongadas usando taxas de variação em cadeia até 2013T2⁹. Em alguns casos, as séries foram prolongadas também para períodos anteriores.

Foram implementadas algumas transformações nas variáveis fornecidas. Em primeiro lugar, calcularam-se alguns rácios, como o crédito sobre o PIB e ativos totais do sistema bancário em percentagem do PIB. Posteriormente foram calculadas taxas de variação homólogas para a maioria das variáveis. Finalmente, foram estimados desvios face às tendências de longo prazo das variáveis usando um filtro *Hodrick-Prescott* unilateral para vários parâmetros de alisamento¹⁰.

Após estas transformações, obtiveram-se 37 possíveis variáveis explicativas. De modo a selecionar as mais relevantes, procedeu-se a uma análise univariada semelhante à descrita por Bonfim e Monteiro (2013), avaliando a AUROC de cada série. Adicionalmente, a disponibilidade dos dados foi também tida em consideração e as séries mais curtas foram excluídas da análise. As variáveis com melhores resultados foram o índice de preços de ações, a taxa de variação homóloga do rácio de serviço da dívida, o desvio do rácio entre crédito e PIB face à sua tendência de longo prazo calculada com um parâmetro de alisamento de 400000¹¹ e a taxa de variação homóloga do índice de preços de imobiliário.

O quadro 1 apresenta algumas estatísticas descritivas destas variáveis para a amostra completa, enquanto o quadro 2 as mostra separadas por país. Para alguns países não existe na base de dados informação sobre algumas das variáveis selecionadas, sendo estes países excluídos da análise multivariada (Bélgica, Bulgária, Chipre, Estónia, Croácia, Hungria, Lituânia, Luxemburgo, Letónia, Malta, Polónia, Roménia, Eslovénia e Eslováquia). A amostra final consiste numa base de dados de 14 países europeus.

Alguns resultados apresentados no quadro 2 merecem ser destacados. Os preços de ações atingiram valores mais elevados em Portugal e França enquanto na Finlândia, Espanha e Suécia registaram valores mais baixos. A taxa de variação homóloga do rácio do serviço da dívida mais alta foi observada na Grécia e no Reino Unido, enquanto na Alemanha e na Finlândia esta variável não se alterou significativamente durante a maior parte do período considerado na amostra. O *gap* do rácio entre crédito e PIB, considerado um dos melhores indicadores de previsão de crises bancárias (Drehmann *et al.*, 2010), apresenta medianas baixas na Alemanha, Países Baixos e Áustria. As medianas mais altas para este *gap* observam-se em Portugal, Irlanda e Itália. Finalmente, os

preços de imobiliário cresceram mais significativamente na Grécia, Reino Unido, Espanha, Irlanda e Finlândia. A dinâmica dos preços de imobiliário apresenta menor magnitude em países como Alemanha, Áustria, França e Portugal.

2.2. Metodologia

Desde o trabalho seminal de Estrella e Hardouvelis (1991) que os modelos de variável binária têm desempenhado um papel importante na estimação e previsão de recessões/cries (Wright, 2006, Kauppi e Saikonen de 2008, e Nyberg, 2009).

Neste artigo consideram-se variantes do modelo *probit* dinâmico geral dado por,

$$y_{it}^* = \alpha + \sum_{k=1}^p \sum_{j=1}^d \beta_{kj} x_{ij,t-k} + \sum_{k=1}^p \gamma_k y_{i,t-k} + u_{it} \quad (1)$$

onde y_{it} é uma variável binária de crises, y_{it}^* é uma variável latente tal que $y_{it} = 1$ se $y_{it}^* > 0$ e 0 caso contrário; $x_{ij,t}$, $j = 1, \dots, p$ corresponde a um conjunto de p variáveis exógenas, e $y_{i,t-k}$, $k=1, \dots, p$, corresponde ao k -ésimo desfasamento temporal da variável de crise.

Tendo por base o modelo em (1), na análise empírica irão considerar-se dois modelos distintos: i) o modelo marginal que resulta de considerar $\gamma_1 = \dots = \gamma_p = 0$ em (1), *i.e.*, só são considerados os impactos das variáveis exógenas nas estimativas das probabilidades de ocorrência de crises e a autocorrelação é tratada como um elemento perturbador que é captado por parâmetros de associação; e ii) o modelo de transição que incorpora na regressão de y_{it}^* explicitamente o comportamento passado da variável de crise (modelo (1) completo). Deste modo, o passado específico de cada unidade pode ser utilizado para a previsão correspondente dessa unidade, ao contrário dos modelos marginais, cujas previsões só consideram os valores das variáveis exógenas.

A estimação destes modelos é realizada pelo método da máxima verossimilhança (MLE). A maximização da função de verossimilhança é um problema altamente não linear, mas que pode facilmente ser resolvido por recurso a métodos numéricos convencionais. De Jong e Woutersen (2011) mostram, no contexto de séries temporais univariadas, que sob condições adequadas de regularidade, a teoria de amostras grandes aplica-se ao estimador MLE do vetor de parâmetros da regressão.

Quadro 1 • Estatísticas descritivas

	Amostra completa					
	N	Média	Desvio Padrão	Min.	Mediana (p50)	Máx.
Variável de crises	4816	0.10	0.29	0	0	1
Índice de preços de ações	2678	58.5	44.2	1	48.0	265.1
Rácio de serviço da dívida (taxa de variação anual %)	2827	0.03	0.11	-0.63	0.02	1.24
Gap do rácio entre crédito e PIB	2285	4.6	12.0	-47.2	2.4	62.5
Índice de preços de imobiliário (taxa de variação anual %)	2834	0.11	0.44	-0.42	0.06	14.42

Fontes: Babecky *et al.* (2012), BCE, BIS, Detken *et al.* (2014), Eurostat, FMI, OCDE, Thomson Reuters e cálculos dos autores.

Nota: A variável de crises toma o valor 1 durante crises bancárias ou períodos de elevada vulnerabilidade em que uma crise possa estar eminente. O índice de preços de ações combina dados do Eurostat com dados do FMI de forma a obter uma série o mais longa possível. A série do rácio de serviço da dívida é proveniente de dados do BCE, calculados seguindo a metodologia de Drehmann e Juselius (2012). O desvio do rácio de crédito sobre o PIB face à sua tendência de longo prazo (*gap* do rácio entre crédito e PIB) foi calculado usando um filtro Hodrick-Prescott recursivo, com o parâmetro de alisamento de 400000. O índice de preços de imobiliário tem como fontes uma combinação de dados do BIS e da OCDE.

Quadro 2 • Estatísticas descritivas por país

	Índice de preços de ações		Rácio de serviço da dívida		Gap do rácio entre crédito e PIB		Índice de preços de imobiliário	
	N	mediana	N	mediana	N	mediana	N	mediana
Áustria	177	33.5	167	0.02	169	0.5	132	0.03
Bélgica	0	.	127	0.02	129	6.9	161	0.06
Bulgária	59	49.8	60	0.15	0	.	132	0.03
Chipre	0	.	75	0.04	0	.	0	.
República Checa	84	59.1	66	0.01	80	4.4	80	0.04
Alemanha	177	54.4	166	0.00	169	-1.4	132	0.03
Dinamarca	98	68.8	35	0.02	141	2.9	132	0.05
Estónia	66	54.6	61	0.10	0	.	30	0.07
Espanha	177	26.6	167	0.01	169	1.4	132	0.07
Finlândia	177	24.0	167	0.00	169	2.4	132	0.06
França	105	79.0	167	0.01	169	1.9	132	0.04
Grécia	85	63.3	41	0.10	129	1.2	132	0.12
Croácia	61	89.1	3	-0.06	0	.	62	0.03
Hungria	93	43.8	67	0.05	0	.	88	0.09
Irlanda	177	32.3	127	0.01	129	8.0	132	0.06
Itália	177	39.7	167	0.01	168	6.1	132	0.05
Lituânia	48	77.4	68	0.14	0	.	55	0.06
Luxemburgo	77	59.6	126	-0.01	0	.	141	0.09
Letónia	67	65.9	68	0.20	0	.	26	0.03
Malta	0	.	160	0.03	0	.	0	.
Países Baixos	177	35.1	166	0.01	141	0.1	149	0.04
Polónia	0	.	27	0.08	84	4.1	96	0.09
Portugal	105	79.7	139	0.01	141	12.5	130	0.04
Roménia	0	.	20	-0.04	0	.	89	0.49
Suécia	177	25.0	167	0.01	129	2.4	132	0.05
Eslovénia	62	55.2	32	0.08	0	.	22	-0.01
Eslováquia	75	44.7	24	0.04	0	.	80	0.07
Reino Unido	177	46.6	167	0.04	169	2.5	173	0.09
Total	2678	48.0	2827	0.02	2285	2.4	2834	0.06

Fontes: Babecky *et al.* (2012), BCE, BIS, Detken *et al.* (2014), Eurostat, FMI, OCDE, Thomson Reuters e cálculos dos autores.
Nota: Todas as variáveis são definidas no quadro 1.

3. Resultados

3.1. Resultados principais

O primeiro passo da análise consiste na estimação dos modelos anteriormente indicados. Assim, denotando a variável binária de crises por y_{it} (que toma o valor 1 se se observar uma crise e 0 caso contrário), podem obter-se projeções a vários horizontes recorrendo a especificações de modelos de dados de painel (*pooled*), onde a previsão da probabilidade de se observar uma crise no período t , $P(y_{it} = 1)$, é dada por $\Phi(y_{it}^*)$. Em particular, $\Phi(\cdot)$ é uma função de distribuição cumulativa Gaussiana e y_{it}^* é portanto uma variável latente. Considerando h como o horizonte de previsão, podemos ajustar (1) para produzir as previsões necessárias, *i.e.*,

$$y_{it}^* = \alpha + \sum_{k=1}^p \sum_{j=1}^d \beta_{kj} x_{ij,t-k-h} + \sum_{k=1}^p \gamma_k y_{i,t-k-h} + v_{it} \quad (2)$$

O modelo em (2) foi estimado com três estruturas de desfasamentos distintas, conforme discutido anteriormente. Primeiro consideraram-se desfasamentos de 4 a 20 trimestres para todas as variáveis explicativas. Isto permite analisar os determinantes de crises bancárias com antecipação de 1 a 5 anos. Adicionalmente estimaram-se os modelos no designado período distante,

explorando os determinantes de crises utilizando desfasamentos entre 12 e 20 trimestres, o que permite detetar as variáveis com sinais de alerta antecipado mais fortes. Finalmente, estimaram-se também os modelos para um período próximo, utilizando informação desfasada de 4 a 12 trimestres, analisando assim quais as variáveis mais relevantes na sinalização de uma crise no futuro próximo.

Para todos os modelos, começou-se por utilizar todos os desfasamentos temporais de todas as variáveis explicativas selecionadas (índice de preços de ações, a variação homóloga do rácio de serviço da dívida, o *gap* do rácio entre crédito e PIB, e a taxa de variação homóloga do índice de preços de imobiliário). Destas estimações selecionaram-se unicamente as variáveis que se mostraram estatisticamente significativas a um nível de significância de 10%, o que permitiu a estimação de modelos mais parcimoniosos, conforme resultados apresentados no quadro 3.

Os resultados relativos ao índice de preços de ações não são particularmente fortes. Na representação parcimoniosa (quadro 3), o índice de preços de ações mostra-se estatisticamente significativo (10%) a $t-6$, $t-9$ e $t-10$ trimestres. O crescimento do rácio de serviço da dívida apresenta sinais de antecipação significativos (em $t-16$, $t-17$ e $t-20$), confirmando deste modo os resultados de Drehmann e Juselius (2012). O *gap* do rácio entre crédito e PIB é a variável que apresenta maior número de coeficientes estatisticamente significativos, com sinais de alerta em muitos trimestres anteriores à crise. No entanto, os sinais destes coeficientes nem sempre são consistentes, *i.e.*, em alguns trimestres os coeficientes estimados são positivos e em outros negativos. Finalmente, a taxa de variação homóloga do índice de preços de imobiliário também apresenta sinais mistos, com coeficientes positivos em $t-5$ e $t-11$, e coeficientes negativos em $t-4$. Isto poderá sugerir que crises bancárias sistémicas são mais prováveis após períodos de forte crescimento dos preços do imobiliário seguidos de um acentuado declínio.

No período distante ($t-12$ a $t-20$), os resultados são algo diferentes. A taxa de variação homóloga do índice de preços de imobiliário não se apresenta estatisticamente significativa, indicando desta forma que esta variável não apresenta propriedades de alerta antecipado adequadas num contexto multivariado. O crescimento do rácio de serviço da dívida é significativo em $t-16$ e $t-17$, mantendo os sinais positivos do período total de estimação. O índice de preços de ações apresenta sinais positivos significativos no desfasamento a 12 trimestres. O *gap* do rácio entre crédito e PIB também é significativo em vários períodos.

No período mais próximo da crise ($t-4$ a $t-12$), a taxa de crescimento do rácio de serviço da dívida nunca é estatisticamente significativa. Isto indica que esta variável tem propriedades fortes de sinalização antecipada de alerta, contudo não próximo da emergência de uma crise. As outras três variáveis continuam a apresentar significância.

Na segunda parte do quadro 3 apresentam-se os resultados para os modelos dinâmicos. Conforme discutido anteriormente, ao explorar a dependência temporal contida na variável de crises, procura-se melhorar a qualidade dos modelos de alerta avançado. De facto, a inclusão de desfasamentos temporais da variável de crises na especificação do modelo parece melhorar substancialmente o ajustamento do mesmo. Vários desfasamentos da variável de crises mostram-se estatisticamente significativos na explicação da probabilidade da ocorrência de uma crise bancária sistémica, nas três janelas de estimação consideradas. Os resultados respeitantes às outras variáveis explicativas são consistentes. A principal exceção é o crescimento do rácio de serviço da dívida, que agora se apresenta significativo também no período próximo.

De um modo geral, o crescimento do rácio de serviço da dívida aparenta dar com elevada antecipação orientações sobre crises ao decisor. O *gap* do rácio entre crédito e PIB dá sinais fortes em todos os horizontes, ainda que nem sempre consistentes.

Quadro 3 • Resultados das regressões: *probits* simples e dinâmicos

	Probit simples						Probit dinâmico											
	Período completo		Período distante [12;20] lags		Período próximo [4;12] lags		Período completo		Período distante [12;20] lags		Período próximo [4;12] lags							
	Lags	Coef.	P> z	Lags	Coef.	P> z	Lags	Coef.	P> z	Lags	Coef.	P> z						
Variável de crises							L4.	6.66	0.00	L12.	1.33	0.00	L4.	7.44	0.00			
							L5.	-4.09	0.00	L13.	-0.20	0.27	L5.	-4.66	0.00			
							L9.	-0.02	0.98									
							L10.	-0.52	0.22									
							L13.	0.44	0.07									
							L18.	-0.90	0.08									
Índice de preços de ações																		
	L5.	0.00	0.44	L12.	0.01	0.06	L5.	-0.01	0.00	L6.	0.02	0.01	L12.	0.01	0.01	L6.	0.01	0.05
	L6.	0.01	0.00	L14.	0.00	0.44	L6.	0.01	0.00	L9.	0.00	0.33	L13.	0.00	0.10	L9.	-0.01	0.22
	L9.	-0.01	0.04	L16.	0.00	0.88	L7.	0.00	0.91	L11.	-0.01	0.02	L16.	0.00	0.34			
	L10.	0.01	0.00	L18.	0.00	0.49	L10.	0.01	0.17	L12.	0.01	0.00	L18.	0.00	0.40			
	L12.	0.00	0.14							L14.	-0.01	0.05						
	L16.	0.00	0.66							L16.	-0.01	0.08						
	L18.	0.00	0.52							L18.	0.01	0.11						
Rácio de serviço da dívida																		
	L9.	0.73	0.66	L16.	4.76	0.00	L5.	9.10	0.00	L12.	5.59	0.00	L4.	4.47	0.01			
	L11.	0.88	0.63	L17.	2.18	0.00	L6.	-5.33	0.08	L16.	3.60	0.02	L12.	4.48	0.00			
	L16.	8.62	0.00	L20.	2.27	0.11	L9.	4.29	0.02	L17.	2.80	0.11						
	L17.	2.14	0.00				L10.	-0.64	0.73									
	L20.	4.80	0.01				L12.	2.11	0.34									
							L13.	2.65	0.19									
							L16.	3.87	0.02									
							L20.	3.40	0.02									
Gap do rácio entre crédito e PIB																		
	L4.	-0.03	0.24	L12.	0.06	0.00	L4.	-0.05	0.06	L5.	-0.03	0.29	L14.	0.01	0.78	L12.	0.00	0.63
	L6.	0.03	0.12	L14.	-0.05	0.01	L6.	0.05	0.00	L6.	0.05	0.07	L17.	-0.04	0.00			
	L10.	0.05	0.00	L17.	-0.06	0.00	L11.	0.06	0.00	L14.	-0.06	0.01	L20.	0.05	0.00			
	L11.	0.06	0.02	L20.	0.06	0.00	L12.	-0.03	0.03	L20.	0.04	0.03						
	L14.	-0.07	0.00															
	L15.	-0.02	0.16															
	L16.	-0.08	0.00															
	L17.	-0.02	0.14															
	L20.	0.10	0.00															
Índice de preços de imobiliário																		
	L4.	-15.42	0.00				L4.	-12.00	0.00	L4.	-6.37	0.01				L4.	-11.21	0.00
	L5.	2.72	0.07				L5.	6.88	0.01	L16.	-0.54	0.74				L5.	9.35	0.00
	L6.	1.55	0.25							L19.	1.46	0.29						
	L11.	3.27	0.00															
Constante	-1.83	0.00	-1.57	0.00	-1.10	0.00	-2.38	0.00	-1.79	0.00	-2.32	0.00						
R2		0.4084		0.1919		0.2275		0.6145		0.2549		0.5723						
N		1316		1471		1521		1274		1480		1417						

Fontes: Babecky *et al.* (2012), BCE, BIS, Detken *et al.* (2014), Eurostat, FMI, OCDE, Thomson Reuters e cálculos dos autores.

Nota: Todas as variáveis são definidas no quadro 1. O período completo refere-se a desfasamentos de [4;20] trimestres, o período distante e o período próximo correspondem a desfasamentos de [12;20] e [4;12] trimestres, respetivamente. Desvios padrão *clustered* por país.

3.2. Avaliação dos modelos

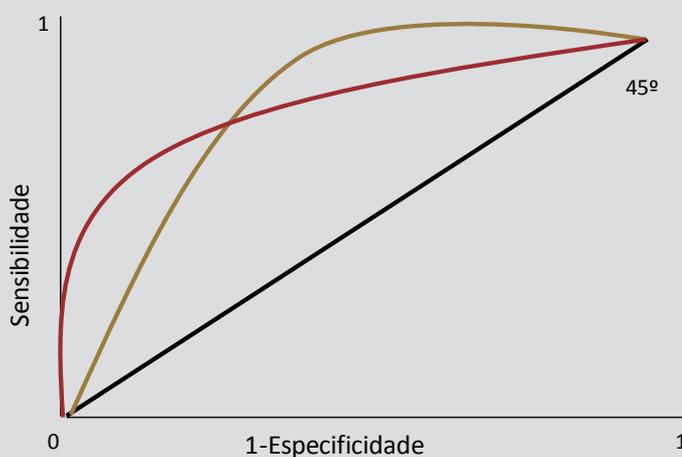
O principal objetivo deste exercício é fornecer uma ferramenta de deteção antecipada de crises bancárias sistémicas que seja útil para decisores de política. Para avaliar a qualidade da orientação dada pelos modelos serão consideradas diversas métricas.

Dado que se trata de um modelo de resposta binária, pode-se definir um limiar para a variável latente. A observação será classificada como “crise” se a variável latente estiver acima desse limiar; caso contrário, a observação é uma “não crise”. Este procedimento define, para cada limiar, uma classificação para cada observação da amostra. Note-se que se sabe, a partir dos dados, a classificação real de cada observação, isto é, aquilo que de facto aconteceu para cada par país-trimestre da amostra. Neste contexto, pode ser construída uma matriz de contingência com quatro elementos: número de verdadeiros positivos (VP, número de observações de crise previstas corretamente pelo modelo), número de verdadeiros negativos (VN, número de observações de não crise previstas corretamente pelo modelo), número de falsos positivos (FP) e número de falsos negativos (FN).

Naturalmente, um modelo perfeito classificaria corretamente todas as observações. Na prática, isto nunca acontece. De facto, um valor muito negativo para o limiar significa que muitas observações de não crise serão classificadas pelo modelo como crises (este é o chamado erro tipo I, e pode-se pensar nele como um falso alarme). À medida que se aumenta o limiar, cada vez mais observações de não crise serão corretamente classificadas pelo modelo, mas algumas observações que realmente correspondiam a crises serão classificadas como não crises (este é o erro tipo II, ou “lobo em pele de cordeiro”). Quando o limiar é muito elevado, todas as observações são classificadas como não crise – de forma que todas as observações de crise serão erradamente classificadas pelo modelo como não crise.

Chama-se **especificidade** à fração de observações de não crise classificadas corretamente pelo modelo, e **sensibilidade** à fração de observações de crise classificadas corretamente pelo modelo. Quando o limiar é menos infinito, todas as observações são classificadas como crise pelo modelo; por conseguinte, a sensibilidade é 1 e a especificidade é 0. Quando o limiar é mais infinito, a sensibilidade é 0 e a especificidade é 1. Fazendo variar o limiar obtemos um conjunto de valores para estas duas medidas. Uma representação possível do desempenho do modelo é a curva *Receiver Operating Characteristic* (ou ROC) que podemos observar no gráfico 1. Este gráfico

Gráfico 1 •
Exemplo de
curvas ROC



Fonte: Cálculos do autor.

ilustra duas curvas ROC hipotéticas. No eixo horizontal representa-se 1 menos a especificidade, isto é, a percentagem de observações de não crise classificadas incorretamente como crise pelo modelo (ou seja, o erro tipo I). No eixo vertical representa-se a sensibilidade, ou seja, a fração de observações de crise classificadas corretamente como crise pelo modelo. Um dado ponto (x,y) na curva clarifica a seguinte questão: Que percentagem x haverá de observações de não crise incorretamente classificadas pelo modelo de forma a classificarmos corretamente uma percentagem y de crises? Como mencionado anteriormente, num modelo perfeito conseguiríamos classificar corretamente 100 por cento das observações sem classificar incorretamente nenhum período de não crise (0 por cento). Isto significa que a curva ROC de um modelo perfeito seria o segmento de reta entre os pontos (0,1) e (1,1). Por outro lado, um modelo que classificasse observações de forma aleatória teria uma curva ROC dada pelo segmento de reta entre os pontos (0,0) e (1,1), isto é, uma linha a 45° com o eixo horizontal. Por outras palavras, o modelo classificaria incorretamente 25 por cento das observações de não crise para classificar corretamente 25 por cento das observações de crise. Este facto sugere que uma medida adequada para avaliar o desempenho do modelo é a área sob a curva ROC, grandeza usualmente designada por AUROC (*area under ROC*).

O gráfico 2 representa as curvas ROC para as seis especificações apresentadas no quadro 3, enquanto o quadro 4 apresenta diversos indicadores da qualidade desses modelos. A qualidade do ajustamento (avaliada através do R^2 de McFadden) e a AUROC fornecem resultados consistentes. O desempenho mais elevado é sempre obtido para o período total da estimação. Em contraste, as estimações para o período distante são as que têm os piores resultados. Isto não é surpreendente, visto ser expectável que os sinais sejam mais fortes na iminência da crise do que 3 anos antes. Não obstante, a consideração de um longo período de informação é relevante, visto que a estimação com o período total da amostra tem melhor desempenho do que a do período próximo (12 a 4 trimestres antes da crise).

No que diz respeito à metodologia, o desempenho do modelo, aferido através do R^2 e da AUROC, é substancialmente melhor quando se incluem efeitos dinâmicos usando a variável dependente desfasada. Isto mostra que a exploração da dinâmica da variável dependente ajuda a melhorar o desempenho do modelo em todos os horizontes considerados.

Quadro 4 • Avaliação dos modelos

	<i>Probit simples</i>			<i>Probit dinâmico</i>		
	Período completo	Período distante	Período próximo	Período completo	Período distante	Período próximo
N	1316	1471	1521	1274	1480	1417
R^2	0.408	0.192	0.228	0.615	0.255	0.572
AUROC	0.898	0.792	0.819	0.959	0.834	0.952
Matriz de contingência - Amostra completa						
Verdadeiros positivos (VP)	149	79	96	190	106	204
Falsos positivos (FP)	46	34	42	29	41	37
Falsos negativos (FN)	113	207	200	63	185	78
Verdadeiros negativos (VN)	1008	1151	1183	992	1148	1098
TOTAL	1316	1471	1521	1274	1480	1417
% falsos alarmes	3.5	2.3	2.8	2.3	2.8	2.6
% crises não previstas	8.6	14.1	13.1	4.9	12.5	5.5
% previsões correctas	87.9	83.6	84.1	92.8	84.7	91.9
Sensibilidade (VP/(VP+FN))	56.9	27.6	32.4	75.1	36.4	72.3
Especificidade (VN/(FP+VN))	95.6	97.1	96.6	97.2	96.6	96.7

Fontes: Babecky *et al.* (2012), BCE, BIS, Detken *et al.* (2014), Eurostat, FMI, OCDE, Thomson Reuters e cálculos dos autores.

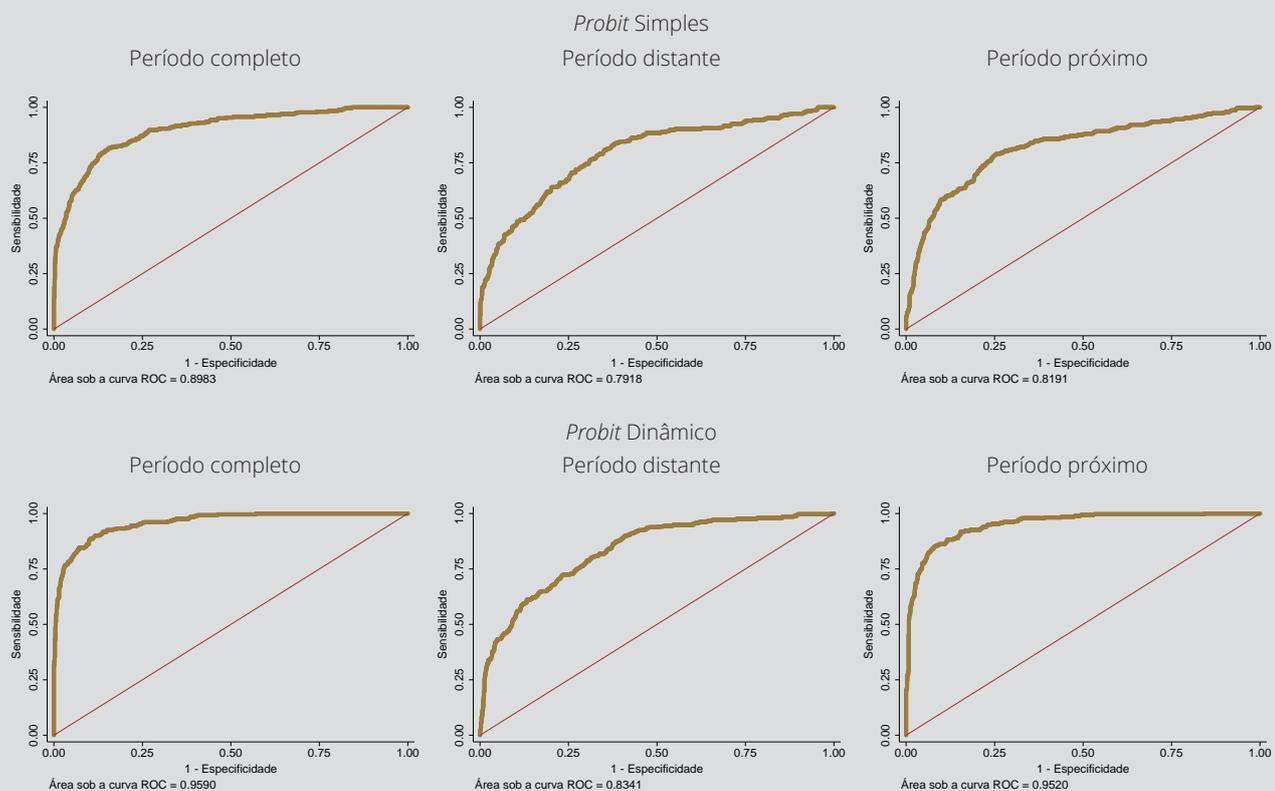
Nota: os resultados referem-se às regressões apresentadas no Quadro 3. Período completo refere-se a desfasamentos de [4;20] trimestres, o período distante e o período próximo correspondem a desfasamentos de [12;20] e [4;12] trimestres, respetivamente.

Embora o ajustamento e o AUROC sejam boas medidas sintéticas para aferir o desempenho de cada modelo, é também relevante considerar quantas crises seria o modelo capaz de prognosticar, quantas vezes falharia e quantos falsos alarmes existiriam. Isto é especialmente relevante no enquadramento presente, com implicações potenciais no que diz respeito à tomada de decisões de política. De facto, como referido por Alessi e Detken (2011), os decisores não são indiferentes entre serem surpreendidos por uma crise ou atuarem sob um falso alarme. Visto que existe um compromisso entre estas duas dimensões, e que está implícito na AUROC, será interessante olhar para ambas separadamente.

Os *probits* dinâmicos permitem reduzir a percentagem de falsos alarmes somente para as estimações no período completo e no período próximo. No entanto, esta percentagem é muito pequena para todos os modelos, sendo no máximo 3.5 por cento (*probit* simples - período completo). Por outro lado, os *probits* dinâmicos reduzem significativamente a percentagem de crises não detetadas (de 8.6 para 4.9 por cento no período de estimação total). Dado que falhar a deteção de uma crise pode ter maiores custos do que ter um falso alarme (Demirgüç-Kunt e Detragiache, 1999, Borio e Lowe, 2002, e Borio e Drehmann, 2009), este resultado sugere que os *probits* dinâmicos podem de facto ser mais úteis para os decisores de política. Para além disso, os *probits* dinâmicos são capazes de prever uma taxa mais elevada de episódios de crise do que os outros modelos considerados, sobretudo no período total e no período tardio.

É também interessante notar que os *probits* dinâmicos permitem aumentar significativamente a sensibilidade. Como mencionado acima, a sensibilidade é a percentagem de positivos verdadeiros (ou seja, de crises corretamente classificadas) sobre o total de ocorrências do evento que

Gráfico 2 • Curvas ROC



se pretende prever (ou seja, de crises). Isto sugere que os *probits* dinâmicos serão mais úteis na identificação de crises do que os modelos marginais. Por outro lado, a especificidade do modelo, definida como a percentagem de negativos corretamente classificados sobre o total de observações em que não se verifica a ocorrência de uma crise, decresce ligeiramente nos modelos dinâmicos, embora permanecendo a níveis muito altos.

Em resumo, esta bateria de métricas confirma que adicionar uma componente dinâmica a modelos de deteção precoce de crises melhora substancialmente a qualidade dos resultados, em especial ao reduzir a percentagem de crises não detetadas e ao aumentar a percentagem daquelas que são corretamente identificadas. Como discutido na secção 2.1, esta metodologia foi parte de uma competição entre diferentes metodologias apresentadas numa conferência do Banco Central Europeu. A esse propósito, Alessi *et al.* (2014) mostram que os *probits* dinâmicos estiveram entre as metodologias com melhor desempenho.

3.3. Robustez

Os resultados apresentados até este momento avaliam a qualidade do modelo na amostra (*in-sample*). Contudo, a qualidade do modelo depende do seu rigor em termos de previsão. Deste modo, é essencial testar o comportamento do modelo dentro e fora da amostra. Para tal, foram considerados dois exercícios distintos. Primeiro, foi implementada uma estimação fora-de-período (*out-of-period*), excluindo o período da crise financeira global da amostra para todos os países (foram excluídos todos os trimestres de 2007T1 em diante). O segundo exercício consistiu numa estimação fora-da-amostra (*out-of-sample*), excluindo a Dinamarca, Finlândia e Suécia, onde ocorreu uma crise bancária sistémica no final dos anos 1980s/início de 1990s, avaliando a qualidade do modelo para estes países (*ex-post*).

Os resultados relativos ao desempenho dos modelos simples e dinâmicos nestes dois exercícios são apresentados no quadro 5. O quadro apresenta vários indicadores de avaliação para os *probits* simples e dinâmicos, para as três janelas de estimação consideradas (período completo, período distante e período próximo). Os resultados *in-sample* são comparados com as estimativas *out-of-period* e *out-of-sample*. Nestes dois casos, os modelos são estimados excluindo, respetivamente, os períodos e os países referidos acima. Os indicadores de avaliação referem-se ao desempenho da previsão do modelo para estas observações excluídas.

Verifica-se que a AUROC para as estimações *out-of-sample* e *out-of-period* não diminui de forma significativa na maior parte das especificações. Pelo contrário, até aumenta nos *probits* simples estimados para o período completo, bem como nas estimações *out-of-sample* dinâmicas. Por seu turno, a AUROC para a estimação *out-of-period* diminui apenas ligeiramente, confirmando deste modo que o desempenho do modelo não depende criticamente da crise financeira global. Esta questão poderia constituir uma preocupação, dado que uma parte substancial das observações em vários países se refere ao período pós-2007.

Contudo, a percentagem de falsos alarmes aumentou ligeiramente, sobretudo nas estimativas *out-of-sample*. Por seu turno, a percentagem de crises não identificadas aumenta de forma mais significativa na estimação *out-of-period*, sugerindo que o modelo não seria capaz de prever a crise financeira global em todos os países da amostra. A percentagem de crises corretamente previstas também diminui mais nesta estimação. Estes últimos resultados não são inesperados, dado que esta crise foi originada em muitos países por choques exógenos e não tanto por vulnerabilidades latentes.

4. Conclusões

As crises bancárias sistémicas são eventos raros, ainda que extremamente onerosos. A sua previsão rigorosa apresenta significativos desafios, apesar da vasta literatura sobre este tema. Neste artigo é apresentada uma contribuição metodológica para esta literatura, explorando o papel de *probits* dinâmicos na previsão destes eventos.

Utilizando um conjunto exaustivo de crises bancárias sistémicas na Europa, concluímos que os preços de ações, o crescimento dos preços de habitação, o *gap* do rácio entre crédito e PIB e o crescimento do rácio de serviço da dívida encontram-se entre os indicadores mais úteis para sinalizar a emergência de crises. Os últimos dois indicadores oferecem os sinais mais fortes e consistentes num contexto multivariado.

Neste artigo mostra-se que a inclusão de uma componente dinâmica na modelização multivariada de crises bancárias sistémicas melhora de forma significativa o desempenho dos modelos. Estes resultados são válidos tanto na amostra, como fora da amostra.

Notas

1. Os autores agradecem aos participantes no *ECB/MaRs Workshop on Early Warning Tools and Tools for Supporting Macroprudential Policies* e num seminário no Banco de Portugal por comentários e sugestões relevantes. As opiniões expressas no artigo são dos autores, não coincidindo necessariamente com as do Banco de Portugal ou do Eurosistema.
2. Banco de Portugal, Departamento de Estudos Económicos e *Nova School of Business and Economics*.
3. Banco de Portugal, Departamento de Estudos Económicos.
4. Banco de Portugal, Departamento de Estudos Económicos.
5. Banco de Portugal, Departamento de Estudos Económicos e *Nova School of Business and Economics*.
6. Apesar destes esforços, a informação não é exatamente a mesma que estaria disponível numa situação real, dado que as variáveis macroeconómicas são sujeitas a revisões *ex-post*. Edge e Meisenzahl (2011) mostram que estas diferenças podem ser substanciais quando se calcula o rácio do crédito sobre o PIB, o que pode afetar medidas de política macroprudencial como o *buffer* de capital contracíclico.
7. Por exemplo, para Portugal, foi adicionado um período de *stress* que não se concretizou numa crise mas no qual se acumularam vulnerabilidades significativas. Com base nisto, os trimestres 1999T1 – 2000T1 foram classificados como um período de *stress*. Para mais detalhes, ver Bonfim e Monteiro (2013).
8. Ver definições detalhadas destes conceitos na secção 3.2 Avaliação dos modelos.
9. A única série utilizada que não foi possível prolongar foi o rácio do serviço da dívida.
10. Para uma representação do impacto de diferentes parâmetros de alisamento num contexto semelhante a este, consultar Bonfim e Monteiro (2013).
11. De acordo com o Comité de Basileia (2010) e Drehmann *et al.* (2010), o desvio do rácio entre crédito e PIB face à sua tendência de longo prazo é o indicador com melhor desempenho na sinalização da necessidade de acumular capital antes de uma crise, quando são analisados vários indicadores para vários países. Assim, o Comité de Basileia (2010) propõe que as decisões relativas ao *buffer* de capital contracíclico sejam ancoradas na magnitude destes desvios (apesar de reconhecer a necessidade de complementar estas decisões com outros indicadores e discricionariedade).

Quadro 5 • Estimação *Out-of-period* e *Out-of-sample*

	Probit simples			Probit Simples <i>Out-of-period</i>			Probit Simples <i>Out-of-sample</i>			Probit Dinâmico			Probit Dinâmico <i>Out-of-period</i>			Probit Dinâmico <i>Out-of-sample</i>			
	Período completo	Período distante	Período próximo	Período completo	Período distante	Período próximo	Período completo	Período distante	Período próximo	Período completo	Período distante	Período próximo	Período completo	Período distante	Período próximo	Período completo	Período distante	Período próximo	
	0.898	0.792	0.819	0.915	0.743	0.790	0.906	0.792	0.806	0.959	0.834	0.952	0.947	0.788	0.943	0.966	0.842	0.953	
Confusion matrix - full sample																			
Verdadeiros positivos	149	79	96	72	12	37	28	8	19	190	106	204	106	16	106	42	20	42	
Falsos positivos	46	34	42	19	0	14	27	15	14	29	41	37	9	6	4	17	22	11	
Falsos negativos	113	207	200	82	142	119	30	50	48	63	185	78	42	140	47	16	40	18	
Verdadeiros negativos	1008	1151	1183	155	174	166	154	197	239	992	1148	1098	153	169	176	156	191	190	
TOTAL	1316	1471	1521	328	328	336	239	270	320	1274	1480	1417	310	331	333	231	273	261	
% falsos alarmes	3.5	2.3	2.8	5.8	0.0	4.2	11.3	5.6	4.4	2.3	2.8	2.6	2.9	1.8	1.2	7.4	8.1	4.2	
% crises não previstas	8.6	14.1	13.1	25.0	43.3	35.4	12.6	18.5	15.0	4.9	12.5	5.5	13.5	42.3	14.1	6.9	14.7	6.9	
% previsões correctas	87.9	83.6	84.1	69.2	56.7	60.4	76.2	75.9	80.6	92.8	84.7	91.9	83.5	55.9	84.7	85.7	77.3	88.9	
Sensibilidade (VP/(VP+FN))	56.9	27.6	32.4	46.8	7.8	23.7	48.3	13.8	28.4	75.1	36.4	72.3	71.6	10.3	69.3	72.4	33.3	70.0	
Especificidade (VN/(FP+VN))	95.6	97.1	96.6	89.1	100.0	92.2	85.1	92.9	94.5	97.2	96.6	96.7	94.4	96.6	97.8	90.2	89.7	94.5	

Fontes: Babcock et al. (2012), BCE, BIS, Deijen et al. (2014), Eurostat, FMI, OCDE, Thomson Reuters e cálculos dos autores.

Nota: Os resultados dos exercícios "out-of-sample" excluem a Dinamarca, Finlândia e Suécia, países que tiveram uma crise bancária no final dos anos 1980/início dos anos 1990. Os resultados "out-of-period" excluem a crise financeira global que começou em 2007. O período completo considera desfasamentos de 4/20 trimestres, enquanto período distante e período próximo correspondem a desfasamentos de [1;2;20] e [4;12] trimestres, respetivamente.

REFERÊNCIAS

- Alessi, L., A. Antunes, J. Babecký, S. Baltussen, M. Behn, D. Bonfim, O. Bush, C. Detken, J. Frost, R. Guimarães, T. Havránek, M. Joy, K. Kauko, J. Matějů, N. Monteiro, B. Neudorfer, T. Peltonen, P. M. M. Rodrigues, M. Rusnák, W. Schudel, M. Sigmund, H. Stremmel, K. Šmídková, R. van Tilburg, B. Vašíček e D. Žigraiová (2014), "A Horse Race of Early Warning Systems Developed by the Macroprudential Research Network", *mimeo*.
- Alessi, L. e C. Detken (2011), "Quasi real time early warning indicators for costly asset price boom/bust cycles: A role for global liquidity", *European Journal of Political Economy*, 27(3), 520-533.
- Babecky, J., Havranek, T., Mateju, J., Rusnak, M., Smidkova, K., e B. Vasicek (2012), "Banking, debt and currency crises: early warning indicators for developed countries", *ECB Working Paper: 1485/2012*.
- Basel Committee (2010), "Guidance for National Authorities Operating the Countercyclical Capital Buffer".
- Bonfim, D. e N. Monteiro (2013), "A implementação do buffer de capital contracíclico: regras versus discricionariedade", *Relatório de Estabilidade Financeira do Banco de Portugal*, Novembro 2013.
- Borio, C. e M. Drehmann (2009), "Assessing the risk of banking crises – revisited", *BIS Quarterly Review*, March, pp 29-46.
- Borio, C. e P. Lowe (2002), "Assessing the risk of banking crises", *BIS Quarterly Review*, pp 43-54.
- Boyd, J., G. De Nicolò e E. Loukoianova (2009), "Banking Crises and Crisis Dating: Theory and Evidence", *IMF Working Paper 09/141*.
- Burnside, C., M. Eichenbaum e S. Rebelo (2004), "Government guarantees and self-fulfilling speculative attacks", *Journal of Economic Theory*, 119(1), 31-63.
- Chang, R. e A. Velasco (2001), "A model of financial crises in emerging markets", *Quarterly Journal of Economics*, 116, 489-517.
- Chaudron, R. e J. de Haan (2014), "Identifying and dating systemic banking crises using incidence and size of bank failures", *DNB Working Paper No. 406*.
- Cecchetti, S., Kohler, M., e C. Upper (2009), "Financial Crises and Economic Activity", *NBER Working Paper 15379*.
- de Jong, R. M., e T. M. Woutersen (2011), "Dynamic Time Series Binary Choice", *Econometric Theory* 27, 673 - 702.
- Demirgüç-Kunt, A. e E. Detragiache (1999), "Monitoring banking sector fragility: A multivariate logit approach with an application to the 1996-97 crisis", *World Bank Policy Research Working Paper No. 2085*.
- Detken, C., Weeken, O., Alessi, L., Bonfim, D., M. Boucinha, M., Castro, C., Frontczak, S., Giordana, G., Giese, J., Jahn, N., Kakes, J., Klaus, B., H. Lang, J., Puzanova, N., e P. Welz (2014), "Operationalising the countercyclical capital buffer: indicator selection, threshold identification and calibration options", *ESRB Occasional Paper*, forthcoming.
- Drehmann, M., Borio, C., Gambacorta, L., Jimenez, G. e C. Trucharte (2010), "Countercyclical Capital Buffers: Exploring options", *BIS Working Paper No. 317*.
- Drehmann, M. e M. Juselius (2012), "Do debt service costs affect macroeconomic and financial stability?", *BIS Quarterly Review*, September, pp. 23-25
- Edge, R. e R. Meisenzahl (2011), "The Unreliability of Credit-to-GDP Ratio Gaps in Real Time: Implications for Countercyclical Capital Buffers", *International Journal of Central Banking*, December 2011, 261-298.

Estrella A, e G.A. Hardouvelis (1991), "The term structure as a predictor of real economic activity", *Journal of Finance* 46, 555-576.

Jordà, O., Schularick, M., e A. M. Taylor (2010), "Financial crises, credit booms, and external imbalances: 140 years of lessons", *NBER Working Paper* 16567.

Jordà, O., Schularick, M., e A. M. Taylor (2012), "When Credit Bites Back: Leverage, Business Cycles, and Crises", *Federal Reserve Bank of San Francisco Working Paper* 2011-27.

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

Kauppi, H. e P. Saikkonen (2008), "Predicting U.S. Recessions with Dynamic Binary Response Models", *Review of Economics and Statistics* 90(4), 777-791.

Krugman, P. (1979), "A Model of Balance-of-Payments Crises", *Journal of Money, Credit and Banking*, 11(3), 311-325.

Nyberg, H. (2010), "Dynamic Probit Models and Financial Variables in Recession Forecasting", *Journal of Forecasting* 29, 215-230

Obstfeld, M. (1986), "Rational and Self-Fulfilling Balance-of-Payments Crises", *American Economic Review*, 76(1), 72-81.

Oet, M., Eiben, R., Gramlich, D., L. Miller, G., e S. J. Ong (2010), "SAFE: An Early Warning System for Systemic Banking Risk", *mimeo*.

Reinhart, C. e R. Rogoff (2011), "This Time Is Different: Eight Centuries of Financial Folly", *Princeton University Press*.

Wright J. H. (2006), "The yield curve and predicting recessions.", *Finance and Economics Discussion Series no. 7*, Board of Governors of the Federal Reserve System.