



Банк России



ИЮЛЬ 2020

Включение индикаторов финансового развития в системы раннего предупреждения

Серия докладов об экономических исследованиях, №58

А. Пономаренко
С. Татаринцев

Алексей Пономаренко

Департамент исследований и прогнозирования, Банк России.

Email: ponomarenkoa@cbr.ru

Стас Татаринцев

Департамент исследований и прогнозирования, Банк России.

Email: tatarintsevs@cbr.ru

Авторы выражают благодарность Рамису Хабибуллину за ценные замечания и предложения.

Серия докладов об экономических исследованиях Банка России проходит процедуру анонимного рецензирования членами Консультативного совета Банка России и внешними рецензентами.

Фото на обложке: Shutterstock.com

© Центральный банк Российской Федерации, 2020

Адрес: 107016, Москва, ул. Неглинная, 12
Телефоны: +7 495 771-91-00, +7 495 621-64-65 (fax)
Сайт: www.cbr.ru

Все права защищены. Настоящий доклад выражает личную позицию авторов, которая может не совпадать с официальной позицией Банка России. Банк России не несет ответственности за содержание доклада. Любое воспроизведение представленных материалов допускается только с разрешения авторов.

Резюме

Мы оценили модель раннего предупреждения финансовых кризисов на основе подхода «случайного леса». Мы используем оригинальный набор прогнозных показателей, включающий индикаторы финансового развития (например, отношения уровней кредита к ВВП) в дополнение к обычным показателям финансового дисбаланса (например, кредитные разрывы). Модель оценивается посредством трехэтапной процедуры (т.е. обучение, валидация и тестирование). Результаты показывают, что сочетание показателей финансовых дисбалансов и финансового развития помогает повысить точность вневыборочного прогноза систем раннего предупреждения.

JEL-классификация: C40, C52, G01, E44

Ключевые слова: индикаторы раннего предупреждения, финансовый кризис, финансовое развитие, кредитный разрыв, «случайный лес»

1. Введение

Последний финансовый кризис стимулировал новую волну исследований моделей раннего предупреждения. Основной проблемой остается заблаговременная реакция на рост рисков в финансовой системе и смягчение последствий надвигающегося финансового кризиса.¹ Отсюда проистекает необходимость создать надлежащую систему раннего предупреждения (СРП) для выбора правильного момента принятия макропруденциальных мер.

Существующие СРП обычно основываются на наборе индикаторов финансовых дисбалансов: в основном на показателе кредитного разрыва (Drehmann et al. 2011), а также на прочих дополнительных циклических показателях финансовых рисков (Giese et al. 2014, Detken et al. 2014, Tölö et al. 2018, Gersl and Jasova 2018). В том случае, если СРП подаёт сигнал об увеличении вероятности финансового кризиса, регулятору стоит рассмотреть возможность введения макропруденциальных мер.

Данная работа вносит вклад в научную литературу в нескольких направлениях. Во-первых, существующие СРП редко используют индикаторы финансового развития (ИФР), хотя их важность для анализа кредитных циклов существенна (Cottarelli et al. 2005, Égert et al. 2006, Buncic and Melecky 2014, Bahadir and Valev 2015, Naceur et al. 2019). Таким образом, в данной работе рассматривается использование показателей ИФР совместно с традиционными показателями финансовых дисбалансов.

Другое нововведение состоит в том, что для анализа результатов мы использовали алгоритмы машинного обучения в духе работ Holopainen and Sarlin 2017, Alessi and Detken 2018, Beutel et al. 2019. А именно, мы применили алгоритм «случайного леса» (Random Forest) для описания нелинейной зависимости между финансовым развитием, финансовыми дисбалансами и вероятностью финансового кризиса.

Материал структурирован следующим образом. В разделе 2 представлен набор индикаторов, используемых в системах раннего предупреждения, и даётся определение финансового кризиса. В разделе 3 представлено описание модели и алгоритм её оценки. В разделе 4 обсуждается структура эксперимента и

¹ См. Jerman and Quadrini (2012), Mumtaz and Zanetti (2016), и Zanetti (2019), показавших критическую важность финансовых факторов для колебаний делового цикла.

представлены основные эмпирические результаты. В разделе 5 более детально описано взаимодействие показателей финансовых дисбалансов и финансового развития. Раздел 6 содержит выводы.

2. Данные

Набор индикаторов раннего предупреждения по категориям представлен в таблице 1. Естественно, в категорию финансовых дисбалансов входят стандартный показатель кредитного разрыва, рекомендованный нормами Базеля III, индикаторы роста реального кредита (Schularick and Taylor 2012, Gersl and Jasova 2018) и темпы роста кредита к ВВП (Jordà et al. 2013). Мы также дополнили набор данных показателем темпа роста ВВП (Borio and Lowe 2002).

Финансовое развитие представлено уровнем отношения кредита к ВВП и индикаторами финансового развития из статистики МВФ (подробное описание см. Svirydzenka 2016). Источниками являются базы данных ОЭСР, БМР и МВФ (подробности см. в таблице 4 в Приложении).

Таблица 1. Список индикаторов

Категория индикаторов	Индикаторы
Кредитный разрыв	Разрыв кредит/ВВП
Финансовые дисбалансы	Темпы роста ВВП за один год Темпы роста ВВП за два года Реальные темпы роста совокупного кредита за один год Реальные темпы роста совокупного кредита за два года Темпы роста кредита к ВВП за один год Темпы роста кредита к ВВП за пять лет
Финансовое развитие	Отношение кредит/ВВП в %% Доступность финансовых организаций (ДФО) Эффективность финансовых организаций (ЭФО) Доступ на финансовые рынки (ДФР)

Все временные ряды имеют квартальную частотность данных, кроме индикаторов доступности и эффективности финансовых организаций и доступа на финансовые рынки, которые имеют годовую частотность. Соответственно, к годовым индикаторам финансового развития мы применяем линейную интерполяцию. В случае отсутствия данных по некоторым отдельным индикаторам мы заполняем отсутствующие результаты наблюдений экстремально большими значениями 10^{18} (означающими, что отсутствие наблюдений никогда не препятствует подаче сигнала). Описание этих индикаторов представлено в Приложении.

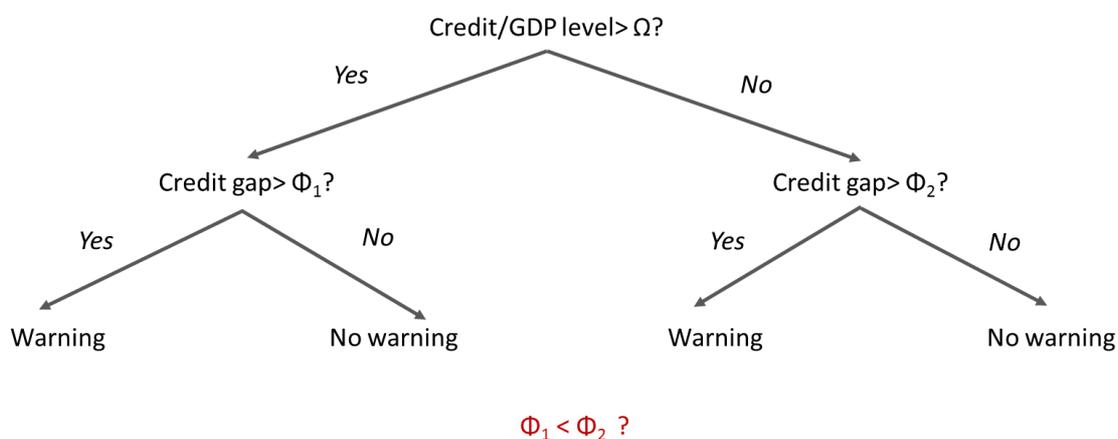
Мы используем базы данных по банковским кризисам, составленные Babecky et al. (2014) и Laeven and Valencia (2018). Первая охватывает эпизоды банковских кризисов, имевших место в 1970–2010 гг., вторая дополняет их данными до 2017 г. В панели мы рассматриваем 30 стран с 1970 Q1 до 2016 Q4, а также дополнительно шесть экономик с формирующимися рынками в период с 2006 Q1 до 2016 Q4 (последние используются только в тестовой подвыборке, как описано в разделе 4). Подробности смотрите в таблице 5 в Приложении. В то же время мы не включаем в основной набор данных результаты наблюдений после 2008 года (хотя результаты для расширенного набора данных отражены в рисунках 6–8 в Приложении). В этот период не было обнаружено новых финансовых кризисов, и полученные посредством этих наблюдений результаты большей частью отражают способность сигнальных индикаторов достаточно быстро угасать, чтобы не допускать ошибочных предупреждений.

Наконец, конструируя зависимую переменную (ряды наличия и отсутствия банковских кризисов), мы учитываем лаг политики и хотим предотвратить предстоящий банковский кризис. Поэтому исходный ряд банковских кризисов преобразован аналогично Alessi and Detken 2018. А именно, 12 кварталов до кризисного события сведены к 1, кризисные события и три квартала до начала кризиса изъяты из набора данных, результаты всех остальных наблюдений сведены к 0. Выборка не сбалансирована: только 13% наблюдений являются кризисными событиями. В результате, мы определяем любой сигнал, появившийся в период за 2–4 года до начала кризиса как правильный сигнал. Сигналы в течение трех кварталов до начала кризиса и в ходе самого кризиса не принимаются во внимание, все остальные сигналы считаются ошибочными.

3. Модель

При построении модели мы базируемся на методе дерева решений. С помощью данного метода мы моделируем связь циклических колебаний финансовых дисбалансов и долгосрочных трендов финансового развития с риском финансового кризиса. Стилизованный пример представлен на рисунке 1. Предполагается, что экономика становится более уязвимой к финансовым кризисам на более высокой степени финансового развития.² Дерево решений может отражать эту связь, распределяя экономики на различные группы исходя из уровня финансового развития на первом этапе и устанавливая различные пороговые значения для переменных финансовых дисбалансов на втором. Это означает, что наш массив данных может включать как циклические, так и нестационарные переменные, что весьма проблематично при использовании альтернативных способов моделирования (например, пробит-регрессии).

Рисунок 1. Пример гипотетического дерева решений



² Это допущение соответствует эмпирическим результатам, демонстрирующим, что превышение финансовым сектором определенного порогового значения негативно влияет на экономический рост (см., например, Cecchetti and Kharroubi 2012, Arcand et al. 2015, Naceur et al. 2018, Yang and Su 2018). Это допущение может также отражать концепцию быстрого роста финансового сектора формирующихся рынков в ходе их догоняющего развития, уже не считающегося благоприятным по достижении конвергенции (см., например, Égert et al. 2006, Buncic and Melecky 2014, Bahadir and Valev 2015). В то же время здесь это допущение используется исключительно в иллюстративных целях. Модель достаточно гибка для учета других типов нелинейного взаимодействия между финансовым развитием и финансовыми дисбалансами.

3.1 Построение модели

«Случайный лес» (Breiman 2001) – популярный метод машинного обучения, широко используемый при решении проблем классификации. Он состоит из десятков или сотен различных деревьев решений. Считается, что это помогает преодолеть переобучение по сравнению с одним бинарным классификационным деревом. Даже если сигналы каждого отдельного дерева могут создавать чрезмерный шум, усреднение сигналов помогает получить менее смещенный и более точный прогноз. Результатом прогнозирования на основании случайного леса является вероятность принадлежности к различным классам (например, к кризисному и некризисному периодам).

Алгоритм использует показатель качества разбиения (Impurity) для разделения подвыборки внутри дерева и выбора пороговых значений для соответствующих индикаторов. Для оценки качества разбиения мы используем индекс Джини:

$$\text{GINI}(f) = \sum_{i=1}^n f_i(1 - f_i)$$

где f_i – доля значений, принадлежащих к классу i .

Мы используем алгоритм «случайного леса» из стандартной библиотеки Python “scikit-learn”. Оптимизация осуществляется для пяти гиперпараметров: четыре определены в функции *RandomForestClassifier* и еще одна описана в функции *predict_proba*. В таблице 2 приведены числовые значения гиперпараметров исходя из которых тестировалась модель.

Таблица 2. Список гиперпараметров

Гиперпараметры	Функция Python	Тестируемые значения
n_estimators	<i>RandomForestClassifier</i>	{20, 30, 40, 50, 60, 70, 80, 90, 100, 300, 800}
max_depth		{2, 3, 5, 7, 10, 12, 15, 18, 25}
max_feature		{None, sqrt, log2}
class_weight		{1, 2, 5, 7, 10, 50, 80, 100, 150, 200, 1500}
Пороговое значение	<i>predict_proba</i>	{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8}

Гиперпараметр $n_estimators$ определяет число деревьев в лесу. Большое число деревьев обеспечивает более точный прогноз. Однако большее число деревьев существенно замедляет процесс обучения. Поэтому мы пытаемся определить оптимальное число деревьев, обеспечивающих баланс между временем обучения и точностью прогноза.

max_depth – другой важный гиперпараметр, представляющий максимальную глубину отдельного дерева. Дерево большой глубины будет более разветвленным и сможет собрать больше информации из данных. С другой стороны, при слишком глубоких деревьях может происходить переобучение (т.е. деревья идеально предсказывают все обучающие данные, но не в состоянии генерализировать результаты для тестовых данных).

$max_feature$ определяет максимальное число индикаторов в отдельном дереве (отражено в таблице 2 как функция от совокупного числа рассматриваемых индикаторов).

$class_weight$ определяет весовые коэффициенты классов для расчета показателя качества разбиения. Этот параметр может быть важен при несбалансированности выборки, что имеет место в данной работе.

Пороговое значение - гиперпараметр, который используется для построения сигнала наступления кризиса, на основании вероятностей, полученных с помощью случайного леса. реализации «случайного леса». При превышения порогового значения мы отмечаем событие как предкризисное: 1 – предкризисное событие, 0 – период без кризиса. Другие гиперпараметры, не определенные в модели, используются со значениями по умолчанию. Полезность (см. раздел 3.2) – показатель оптимизации для выбора лучшей модели и, следовательно, гиперпараметров.

3.2 Оценка модели

Прогнозная способность СРП обычно тестируется с помощью ROC-кривой, однако у этого подхода есть недостатки. Площадь под ROC-кривой даёт обобщённую характеристику качества прогноза, из-за чего она может недооценивать качество прогнозирования в условиях несбалансированности кризисных и некризисных периодов в выборке (см. Baker and Pinsky 2001). Критика ROC-подхода связана с тем, что он придает равный вес ошибкам всех типов, в то время как при прогнозировании финансовых кризисов важность ошибок I и II рода

может быть различна. Если издержки неправильной классификации не одинаковы, суммирование всех возможных пороговых значений ошибочно (Adams and Hand 1999).

Более корректный подход к оценке обеспечивает анализ, основывающийся на показателе “полезности”, рассчитываемый из допущения, что потери властей от одной и той же доли ложных сигналов и неверно предсказанных некризисных периодов могут отличаться. Мы считаем важным протестировать качество индикаторов для разного набора предпочтений. Поэтому (как, например, у Alessi and Detken 2011) оценка модели основывается на индикаторе полезности:

$$usefulness = \min(\theta, 1 - \theta) - loss \quad (1)$$

$$loss = \frac{\theta * FN}{TP + FN} + \frac{(1 - \theta) FP}{TN + FP}, \quad (2)$$

где FN – ложно-отрицательные результаты, FP – ложно-положительные результаты, TN – истинно-отрицательные результаты, TP – истинно-положительные результаты. Полезность показывает улучшение результата СРП по сравнению с наивным подходом в процессе принятия решений. Параметр θ показывает предпочтение центрального банка, большее значение θ указывает на то, что регулятор предпочитает скорее избежать пропуска кризиса, чем подать ложный сигнал. Независимо от результатов прогнозирования СРП лица принимающие решения могут достичь уровень потерь, равный $\min[\theta ; 1 - \theta]$. Если сигнал подаётся во все периоды времени, то $TP = FP = 0$ и значение потерь равно θ . Если же прогнозируются всегда некризисные периоды, то $FN = TN = 0$ и потери становятся равными $(1 - \theta)$. По этой причине (при условии отсутствия результатов СРП) если $\theta < 0.5$, то лицам, принимающим решения выгодно никогда не подавать сигнал, но если $\theta \geq 0.5$, то им выгодно всегда подавать сигнал. В дальнейшем мы будем называть такую стратегию «наивной». Рассматриваемая СРП тестируется на всём диапазоне значений θ .

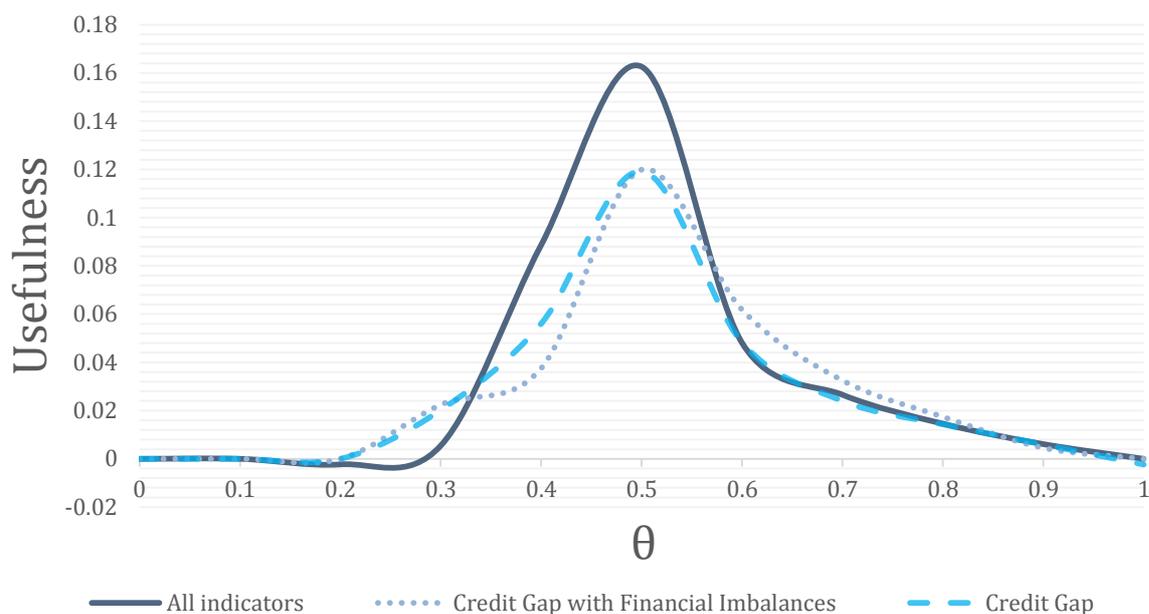
4. Результаты

Как отмечают Beutel et al. (2019), обычный подход двухэтапной валидации, предполагающий разделение результатов наблюдения на две части для оценки и прогнозирования, может не подходить для моделей со значительной гибкостью в построении. А именно, хотя прогнозная (оценочная) выборка не используется для

оценки модели, она тем не менее может использоваться для выбора модели (т.е. если с ее помощью подбираются различные значения гиперпараметров). Соответственно, полезность стратегии в целом может быть переоценена. Поэтому осуществляется более затратная трехэтапная валидация. Мы делим массив данных на три части: “обучающая” подвыборка для обучения моделей, “валидационная” подвыборка для выбора модели и “тестовая” подвыборка для оценки эффективности модели.

Во-первых, мы приводим результаты общепринятой в литературе процедуры “K-fold” кросс-валидации: данные разбиваются на K блоков (см. например Murphy 2012). В нашем случае $K = 5$. Таким образом, данные в выборке разбиваются на пять частей. Мы анализируем модель, используя 3 блока данных за исключением частей i^{th} и j^{th} . Проверяем различные модели (т.е. модели с различными значениями гиперпараметров) на части i^{th} . Далее, выбирается наиболее эффективная модель и тестируется на части j^{th} . Процедура повторяется с использованием всех пяти частей в качестве тестовых подвыборок и рассчитывается общий индикатор полезности для каждой модели. Данная процедура проводится для всего диапазона предпочтений θ . Модели анализируются с использованием либо полного набора объясняющих переменных, либо набора, не включающего ИФР. Также рассматривается поведение отдельного индикатора кредитного разрыва. Результаты моделирования представлены на рисунке 2.

Рисунок 2. Полезность для различных наборов индикаторов (“K-fold” подход)



Как и ожидалось, максимальное значение функции полезности достигается при $\theta = 0,5$. Положительные значения полезности говорят о том, что использование СРП предпочтительнее «наивного подхода». Причём наибольший выигрыш реализуется при $\theta = 0,5$. Снижение полезности при θ , отклоняющемся от 0,5, говорит о том, что СРП и «наивный» подход становятся эквивалентными при росте отклонения θ от значения 0,5. Заметим, что СРП, использующая полный набор индикаторов, выигрывает у конкурентов при большинстве θ ниже 0,6.

Также проведен альтернативный эксперимент, в котором рассмотрена возможность прогнозирования мирового финансового кризиса. Этот подход мы называем «историческим». «Тестовая» подвыборка включает результаты наблюдений 2004–2007 гг. Также в «тестовую» подвыборку добавлены данные по некоторым странам с формирующимися рынками (см. таблицу 4 в Приложении).³

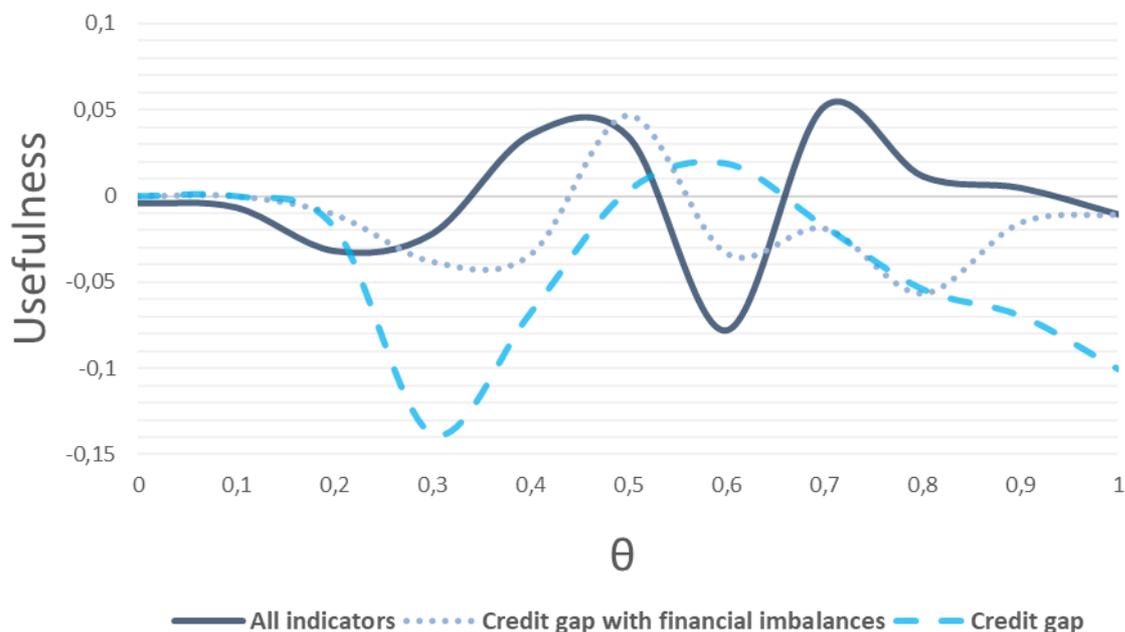
Остаток выборки разделен на пять частей путем создания пяти случайных выборок с равным соотношением числа кризисных и некризисных периодов наблюдений в каждой части. Аналогично подходу «K-fold» происходит оценка модели с использованием подвыборок, состоящих из четырех частей («обучающая» подвыборка). Модель проверяется по оставшейся пятой части («валидационная подвыборка»). Процедура повторяется пять раз с использованием для валидации каждой части, результаты (а именно FP, FN, TP и TN) суммируются, после чего рассчитывается показатель полезности. Данная процедура повторяется для различных наборов гиперпараметров и для каждого θ выбирается спецификация, дающая большую полезность. В заключение, эффективность выбранных моделей оценивается с использованием «тестовой» подвыборки. Результаты представлены на рисунке 3. Результаты для тестовой выборки, включающей наблюдения за 2004–2016 гг., представлены на рисунке 6 в Приложении. Детальные результаты (отдельно для FP, FN, TP и TN по странам) представлены на рисунках 7 и 8 в Приложении.

Полезность всех СРП, выявленная в ходе «исторического» эксперимента, значительно ниже по сравнению с результатами подхода «K-fold» из первого эксперимента. Более того, полезность СРП отрицательна во многих случаях. Снижение эффективности СРП может объясняться характеристиками тестовой

³ На практике, по причине ограниченности данных, применение СРП, оцененных для развитых экономик, к формирующимся рынкам, встречается не так уж редко (см., например, Popomarenko 2013). Исключение этих наблюдений из массива данных не ведет к радикальному изменению результатов.

выборки: данные этой выборки не использовались ни на обучении моделей, ни на их валидации. Также, полученные результаты демонстрируют, что подход «случайного леса» по эффективности превосходит отдельный индикатор кредитного разрыва для большинства параметров θ . Добавление ИФР к СРП однозначно улучшает возможность в прогнозировании мирового финансового кризиса в тех случаях, когда лица принимающие решения предпочитают предотвращать наступление кризисов (При большом значении θ).

Рисунок 3. Полезность на основе различных наборов индикаторов («исторический» подход)



5. Роль индикаторов финансового развития

В данном разделе детальной рассматривается взаимодействие индикаторов финансового развития, финансовых дисбалансов и рисков финансового кризиса. С целью иллюстрации этого взаимодействия проведена следующая процедура. Мы оценили СРП (для $\theta=0,5$) на основе полного набора данных с использованием лучшей спецификации, соответствующей «историческому» подходу. В таблице 3 показано значение важности по Джини (feature importance, подробности см. Menze et al. 2009) отдельных индикаторов в СРП.

Значение важности по Джини показывает, что все индикаторы используются в СРП. Как и ожидалось, значимую роль играет отношение кредита к ВВП (в разрывах, темпах роста и уровнях), хотя задействованы и другие вспомогательные

индикаторы. Далее рассматривается взаимодействие между кредитным разрывом и показателем кредита к ВВП и другими индикаторами ИФР.

Мы рассчитываем возможность кризиса для 99 семплированных массивов данных. Значения ИФР в каждом из этих массивов данных зафиксированы на уровне $i = \{1...99\}$ перцентиля соответствующего распределения⁴, а значения финансовых дисбалансов случайно выбраны из эмпирического распределения (100000 выборок сделаны для каждого i). Полученное распределение представлено на тепловой карте (рисунок 4) и отражает плотность вероятности кризиса в зависимости от уровня финансового развития.

Таблица 3. Значение важности по Джини

Индикаторы	Значение важности по Джини
<i>Финансовые дисбалансы</i>	
Разрыв кредит/ВВП	18,58%
Темпы роста ВВП за один год	3,27%
Темпы роста ВВП за два года	4,44%
Темпы роста отношения кредит/ВВП за один год	12,40%
Темпы роста отношения кредит/ВВП за два года	14,91%
Темпы роста совокупного кредита за один год	7,04%
Темпы роста совокупного кредита за два года	10,16%
<i>Финансовое развитие</i>	
Отношение кредита к ВВП	11,31%
Доступность финансовых организаций	5,19%
Эффективность финансовых организаций	9,40%
Доступ на финансовые рынки	3,29%

Результаты показывают, что вероятность кризиса в целом выше на поздних этапах финансового развития. Заметим, что дисперсия плотности вероятности финансового кризиса становится ниже и два ее пика⁵ более выражены при более высоких значениях ИФР (ядерная плотность вероятности кризиса для ИФР на

⁴ При этом во всех случаях большие значения показателей ИФР соответствуют более высокому уровню финансового развития.

⁵ Предположительно эти два пика представляют пороговое значение важного индикатора финансового дисбаланса.

уровнях 25^{го} и 75^{го} перцентиля представлена на рисунке 5). Данный факт показывает, что взаимодействие между финансовыми дисбалансами и значениями ИФР улавливается моделью. Более высокие уровни финансового развития увеличивают вероятность кризиса, при этом оценки становятся более чувствительными к колебаниям индикаторов финансовых дисбалансов.

Рисунок 4. Вероятность финансового кризиса на различных уровнях финансового развития

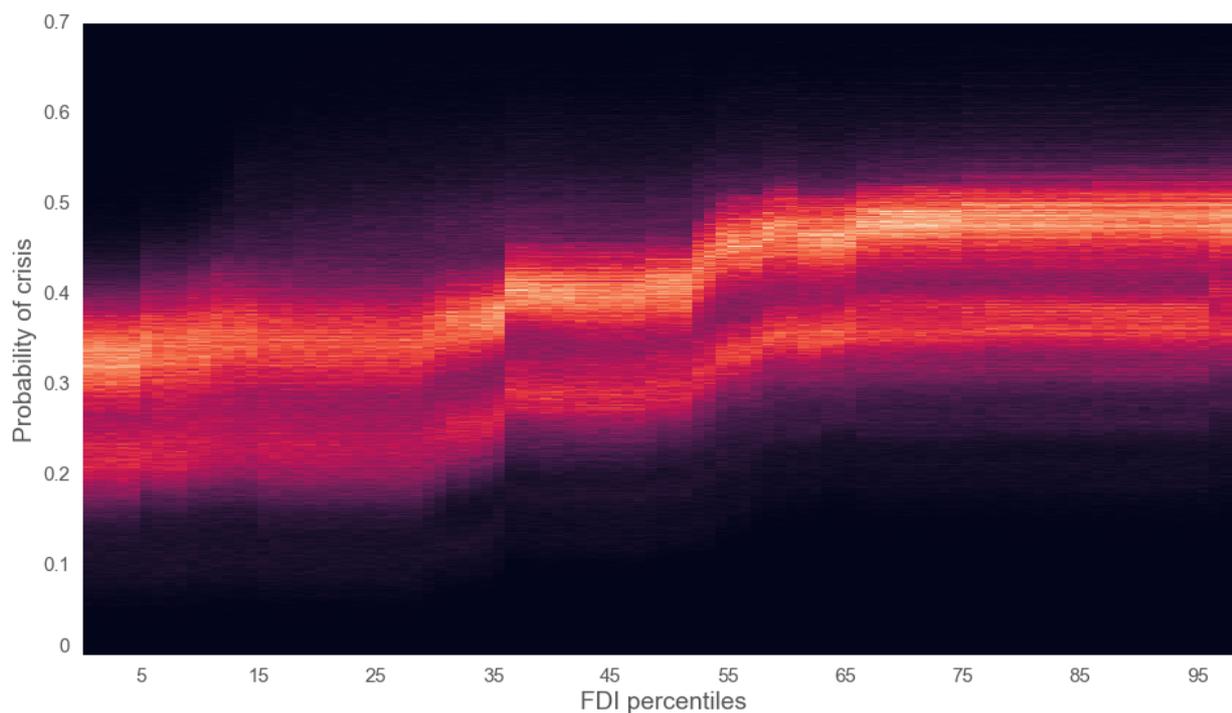
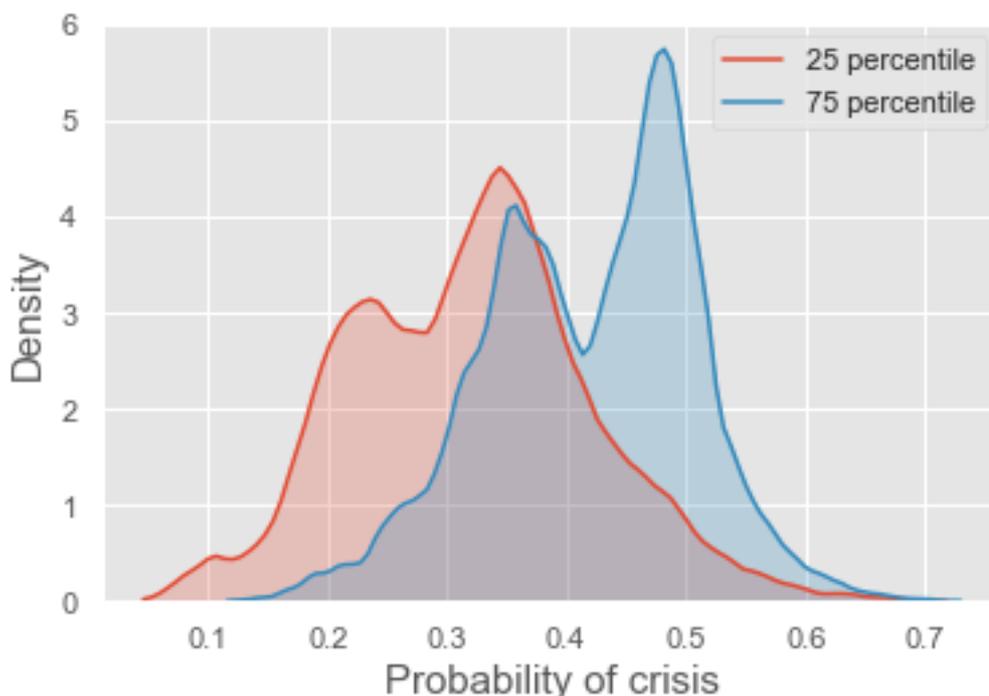


Рисунок 5. Ядерная плотность (%) вероятности финансового кризиса на различных уровнях финансового развития (на 25^{ом} и 75^{ом} перцентилях ИФР)



6. Заключение

Мы создали систему раннего предупреждения финансовых кризисов на основе алгоритма «случайного леса». Мы используем оригинальный набор прогнозных показателей, включающий индикаторы финансового развития (например, отношения уровней кредита к ВВП) в дополнение к обычным показателям финансовых дисбалансов (например, кредитные разрывы). Модель оценивается посредством трехэтапной процедуры (т.е. обучение, валидация и тестирование). Результаты показывают, что сочетание показателей финансовых дисбалансов и финансового развития помогает повысить вневыборочную точность систем раннего предупреждения. Вероятность кризиса в целом выше на поздних этапах финансового развития. Модель также учитывает взаимодействие между финансовыми дисбалансами и ИФР: при более высоких ИФР увеличивается вероятность кризиса и оценки становятся более чувствительными к колебаниям индикаторов финансовых дисбалансов.

Библиография

- 1) Adams, N.M., Hand, D.J. (1999). "Comparing classifiers when the misallocation costs are uncertain." *Pattern Recognition*, 32: 1139–1147.
- 2) Alessi, L., Detken, C. (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: 520–533.
- 3) Alessi, L., Detken, C. (2018). "Identifying excessive credit growth and leverage." *Journal of Financial Stability*, 35: 215–225.
- 4) Arcand, J.L., Berkes, E., Panizza, H. (2015). "Too much finance?" *Journal of Economic Growth*, 20 (2):105–148.
- 5) Babecky, J., Havranek, T., Mateju, J., Rusnak, M., Smidkova, K., Vasicek, B. (2014). "Banking, bedt and currency crisis: stylized facts and early warning indicators." *Journal of Financial Stability*, 15: 1–17.
- 6) Bahadir, B., Valev, N. (2015). "Financial development convergence." *Journal of Banking and Finance*, 56: 61–71.
- 7) Baker, S., Pinsky, P. (2001). "A proposed design and analysis for comparing digital and analog mammography: special receiver operating characteristic methods for cancer screening." *Journal of the American Statistical Association*, 96: 421–428.
- 8) Beutel, J., List, S., Schweinitz, G. (2019). "Does machine learning help us predict banking crises?" *Journal of Financial Stability*, 45, 100693.
- 9) Borio, C., Lowe, P. (2002). "Asset Prices, Financial and Monetary Stability: Exploring the Nexus." BIS Working Paper No. 114.
- 10) Breiman, L. (2001). "Random Forests." *Machine Learning*, 45(1): 5–32.
- 11) Buncic, D., Melecky, M. (2014). "Equilibrium credit: The reference point for macroprudential supervisors." *Journal of Banking & Finance*, 41: 135–154.
- 12) Cecchetti S.G., Kharroubi E. (2012). "Reassessing the impact of finance on growth." BIS Working Papers No. 381.
- 13) Cottarelli, C., Dell’Ariccia, G., Vladkova-Hollar, I. (2005). "Early birds, late risers, and sleeping beauties: Bank credit growth to the private sector in Central and Eastern Europe and in the Balkans." *Journal of Banking and Finance*, 29 (1): 83–104.
- 14) Detken, C., Weeken, O., Alessi, L., Bonfm, D., Boucinha, M., Castro, C., Frontczak, S., Giordana, G., Giese, J., Jahn, N., Kakes, J., Klaus, B., Lang, J., Puzanova, N.,

- Welz, P. (2014). "Operationalising the Countercyclical Capital Buffer: Indicator Selection, Threshold Identification and Calibration Options." ERSB Occasional Paper Series No. 5.
- 15) Drehmann, M., Borio, C., Tsatsaronis, K. (2011). "Anchoring Countercyclical Capital Buffers: The Role of Credit Aggregates." BIS Working Papers No. 355.
- 16) Égert, B., Backé, P., Zumer, T. (2006). "Credit growth in Central and Eastern Europe—new (over)shooting stars?" ECB Working Paper No. 687, European Central Bank.
- 17) Geršl, A. and Jašová, M. (2018). "Credit-based Early Warning Indicators of Banking Crises in Emerging Markets." *Economic Systems*, 42: 18–31.
- 18) Giese, J., Andersen, H., Busha, J., Castro, C., Farag, M., and Kapadia, S. (2014). "The credit-to-gdp gap and complementary indicators for macroprudential policy: evidence from the UK." *International Journal of Finance and Economics*, 19: 25–47.
- 19) Holopainen, M., Sarlin, P. (2017) "Toward robust early-warning models: a horse race, ensembles and model uncertainty." *Quantitative Finance*, 17(12): 1933–1963.
- 20) Jermann, U., Quadrini, V. (2012) "Macroeconomic Effects of Financial Shocks." *American Economic Review* 102(1), 238-71.
- 21) Jordà, Ò., Schularick, M., Taylor, A.M. (2013) "When Credit Bites Back." *Journal of Money, Credit, and Banking*, 45(2): 3–28
- 22) Laeven, L., Valencia, F. (2018). "Systemic Banking Crisis Revised." IMF Working Paper, WP/18/206.
- 23) Menze, B., Kelm, M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., Hamprecht, F. (2009). "A comparison of Random Forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data." *BMC Bioinformatics* 10:213.
- 24) Mumtaz, H., Zanetti, F. (2016). "The Effect Of Labor And Financial Frictions On Aggregate Fluctuations," *Macroeconomic Dynamics*, 20(1), 313-341.
- 25) Murphy, K.P. (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press, Cambridge, Massachusetts.
- 26) Naceur, S.B., Candelon, B., Lajaunie, Q. (2019). "Taming financial development to reduce crises." *Emerging Markets Review*, 40, 100618.
- 27) Ponomarenko, A. (2013). "Early warning indicators of asset price boom/bust cycles in emerging markets." *Emerging Markets Review*, 15(C): 92–106.

-
- 28) Schularick, M., Taylor, A.M. (2012). "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008." *American Economic Review*, 102 (2): 1029–61.
- 29) Sviryndzenka, K. (2016) "Introducing a New Broad-based Index of Financial Development." IMF Working Paper 16–5.
- 30) Tölö, E., Laakkonen, H., Kalatie, S. (2018). "Evaluating Indicators for Use in Setting the Countercyclical Capital Buffer." *International Journal of Central Banking*, 14(2): 51–112.
- 31) Yang, L., Su, J.-J. (2018). "Debt and growth: Is there a constant tipping point?" *Journal of International Money and Finance*, 87: 133–143.
- 32) Zanetti, F. (2019). "Financial Shocks, Job Destruction Shocks, And Labor Market Fluctuations." *Macroeconomic Dynamics*, 23(3), 1137-1165.

Приложение

Таблица 4. Источники данных

Серии	Период	Трансформация	Источник
Кризисные события	1970 Q1:2010 Q4	-	Babecky et al. (2014)
Кризисные события	2010 Q1:2016 Q4	-	Laeven and Valencia (2018)
Реальный ВВП	1970 Q1:2016 Q4	-	Данные ОЭСР
Разрыв кредит/ВВП	1970 Q1:2016 Q1	-	Статистика БМР
Совокупный кредит в местной валюте	1970 Q1:2016 Q1	Лог, дефлятор ВВП	Статистика БМР
Кредит/ВВП, в %%	1970 Q1:2016 Q1	-	Статистика БМР
Доступность финансовых организаций	1980:2016	-	Данные МВФ
Эффективность финансовых организаций	1980:2016	-	Данные МВФ
Доступ на финансовые рынки	1980:2016	-	Данные МВФ

Определения индикаторов финансового развития взяты с сайта МВФ:

Индекс доступности финансовых организаций (ДФО) – компилятивные данные о количестве банковских отделений и банкоматов на 100 тысяч населения;

Индекс эффективности финансовых организаций (ЭФО) – компилятивные данные о чистой процентной марже, спреде между депозитными и кредитными ставками, отношении непроцентного дохода к совокупному доходу, отношении накладных расходов к совокупным активам, доходности активов и доходности капитала банковского сектора;

Индекс доступа на финансовые рынки (ДФР) – компилятивные данные о доле рыночной капитализации, приходящейся на компании за пределами 10 крупнейших, и совокупном числе заемщиков (отечественных и внутренних, нефинансовых и финансовых корпораций) на 100 тыс. взрослого населения.

Таблица 5. Доступные данные по странам

Страны	Периоды наблюдения	Страны	Периоды наблюдения	Страны	Периоды наблюдения
Австралия	1971–2016	Ирландия	1991–2016	Испания	1971–2016
Австрия	1971–2016	Израиль	1996–2016	Швеция	1971–2016
Бельгия	1971–2016	Италия	1971–2016	Швейцария	1971–2016
Канада	1971–2016	Япония	1971–2016	Турция	1999–2016
Чили	1987–2016	Южная Корея	1971–2016	Великобритания	1971–2016
Чешская Республика	1993–2016	Люксембург	2000–2016	США	1971–2016
Дания	1995–2016	Мексика	1981–2016	Бразилия	2006–2016
Финляндия	1971–2016	Нидерланды	1971–2016	Колумбия	2006–2016
Франция	1971–2016	Новая Зеландия	1971–2016	Индия	2006–2016
Германия	1992–2016	Норвегия	1971–2016	Индонезия	2006–2016
Греция	1996–2016	Польша	1996–2016	Российская Федерация	2006–2016
Венгрия	1992–2016	Португалия	1971–2016	ЮАР	2006–2016

Выделенные **жирным** используются только в “исторической” тестовой подвыборке согласно разделу 4. Основной массив данных включает результаты наблюдений только до 2008 г.

Рисунок 6. Полезность на основе различных наборов индикаторов (подход “исторический” подход, тестовая выборка за 2004–2016 гг.)

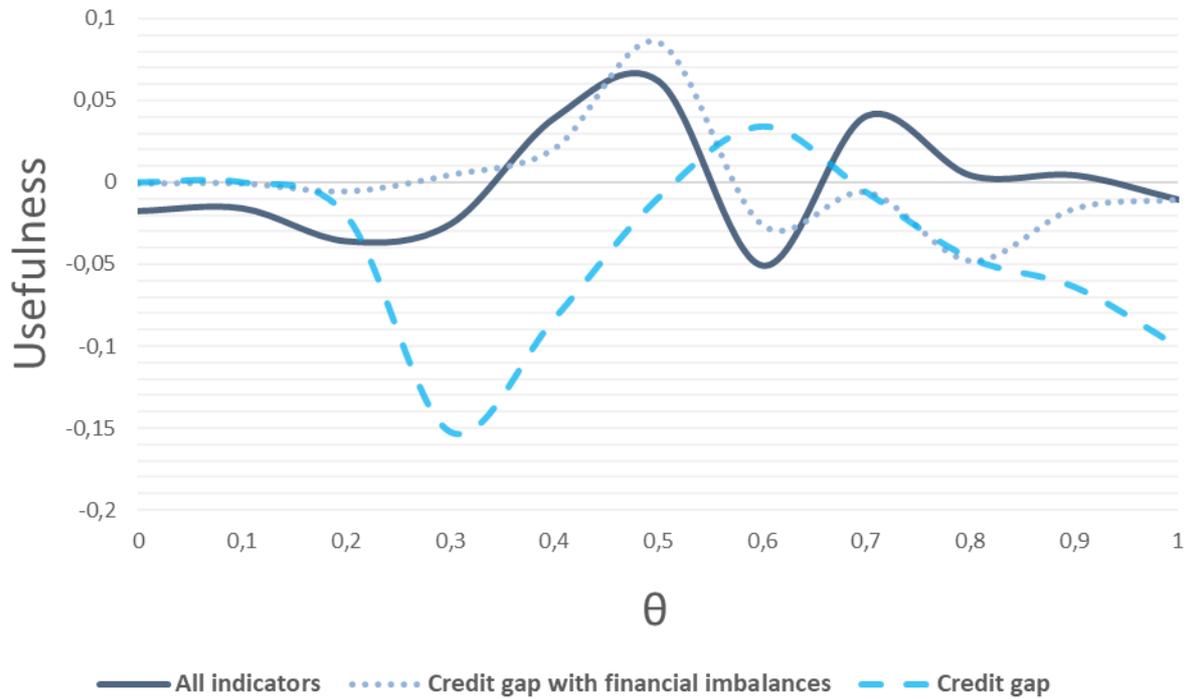


Рисунок 7. Подробные результаты прогнозирования финансовых кризисов в 2004–2016 гг.

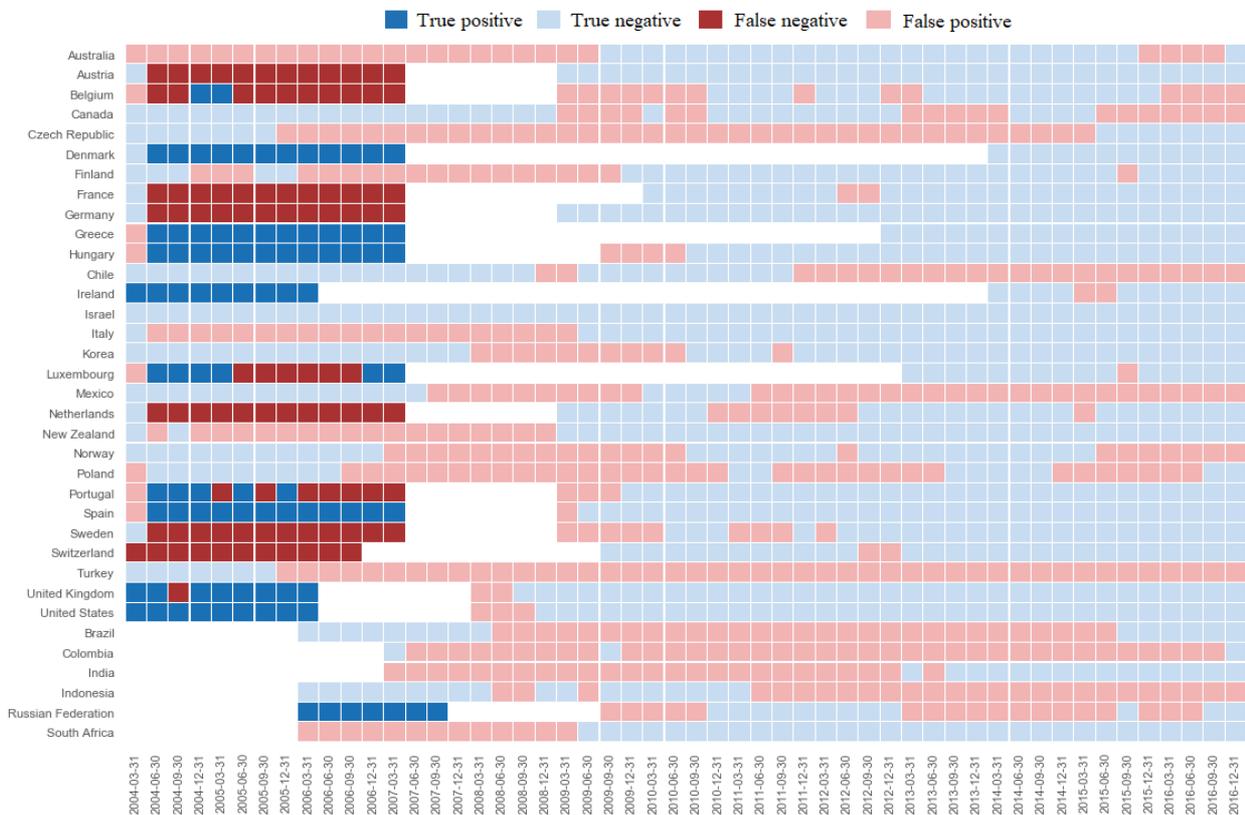


Рисунок 8. Подробные результаты прогнозирования финансовых кризисов в 2004–2016 гг. (доли по каждой стране).

