

---

in the management system development of companies with state participation”, which are as follows: formulated new definitions of the terms in the context of methodology of development of accounting and control processes; the conditions for expanding information boundaries are defined by combining state financial control and internal corporate control, including in the methods of investigating the causes of financial discipline violations, methods of internal corporate control, corporate tax monitoring, and analysis control methods, which is facilitated by the current level of IT technology development and further digitalization of the economy. It is concluded that it is necessary to adopt a law on financial control for national purposes, which is still not available in the country's legal practice. This Law should contain an institutionally recognized and clearly defined term “financial control”, as well as features of control procedures at the national and corporate levels.

*Keywords:* concept, accounting and control processes, management, companies with state participation, internal corporate financial control, digital economy.

*JEL:* M11.

*Manuscript received 01.08.2019*

---

## ОБНАРУЖЕНИЕ ПРИЗНАКОВ ГОРИЗОНТАЛЬНОГО СГОВОРА ПРИ ГОСУДАРСТВЕННЫХ ЗАКУПКАХ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

*Г.О. Молчанова, А.И. Рей,  
Д.Ю. Шагаров*

**DOI:** 10.33293/1609-1442-2020-1(88)-109-127

Совершенствование закупочных процедур и их цифровизация помогают предотвращать и выявлять картели, но в то же время приводят к появлению новых антиконкурентных схем поведения. В нашем исследовании мы концентрируем внимание на электронных аукционах, которые стали в последние годы основным способом осуществления государственных закупок в России. Тот факт, что электронные аукционы дают доступ к большему числу крупных государственных заказов, усиливает стимулы участников торгов заключать антиконкурентные соглашения. Поэтому улучшение методов выявления картелей на электронных аукционах становится особенно актуальной проблемой. Цель данной работы состояла в разработке метода для обнаружения признаков горизонтального сговора на торгах. С помощью методов машинного обучения

---

© Молчанова Г.О., Рей А.И., Шагаров Д.Ю., 2020 г.

*Молчанова Глафира Олеговна*, младший научный сотрудник, Институт отраслевых рынков и инфраструктуры РАНХиГС при Президенте РФ, Москва, Россия; ORCID 0000-0001-8130-1259; molchanova-go@ranepa.ru  
*Рей Алексей Игоревич*, к.э.н., заведующий лабораторией, Институт отраслевых рынков и инфраструктуры РАНХиГС при Президенте РФ, Москва, Россия; ORCID 0000-0001-8207-1790; rey-ai@ranepa.ru  
*Шагаров Дмитрий Юрьевич*, младший научный сотрудник, Институт отраслевых рынков и инфраструктуры РАНХиГС при Президенте РФ, Москва, Россия; ORCID 0000-0001-6544-5359; shagarov-dy@ranepa.ru

мы тренируем классификаторы, которые предсказывают наличие или отсутствие картеля в электронных аукционах в зависимости от распределения ставок участников аукциона, времени подачи заявок, длительности проведения торгов и числа участников. Переменные для модели были отобраны на основании графиков распределений, построенных для выборки по картелям и случайной выборки. Исследование проводится на основе данных с сайта государственных закупок и информации о сговорах на торгах из дел Федеральной антимонопольной службы (ФАС). Результаты показали, что модель случайного леса наиболее точно предсказывает выявление картеля на электронном аукционе. Правильность предсказания – 84%, а полнота и точность модели – 83 и 87%. Наиболее значимыми для классификации переменными оказались уровень снижения цены, разница во времени подачи заявок и начальная (максимальная) цена контракта.

*Ключевые слова:* выявление картелей, сговоры на торгах, государственные закупки, машинное обучение.

*JEL:* C12, C38, C63, D22, D44, H57, L40, L41.

В ходе государственных закупок на рынках нередки ситуации, когда участники вместо того, чтобы конкурировать друг с другом за государственные контракты, распределяют между собой доли на рынке государственного заказа и завышают цены. В целях раскрытия этой противозаконной и наносящей ущерб общественному благосостоянию деятельности могут применяться различные методы. В настоящей работе с использованием методов эконометрики, математической статистики и машинного обучения прогнозируются вероятности того, что закупка будет предметом успешной жалобы в ФАС на картельный сговор между поставщиками.

Современные технологии машинного обучения в состоянии выделить в имеющихся данных по закупкам паттерны, указывающие на сговор между потенциальными поставщиками (горизонтальный сговор) или между заказчиком и одним или более поставщиками (вертикальный сговор). За исключением двух работ, в научной литературе эти методы пока редко использовались для построения

моделей, выявляющих картели (Huber, Imhof, 2019; Ivanov, Nesterov, 2019).

В данном исследовании мы концентрируем внимание на открытых электронных аукционах (ОАЭФ) как на наиболее популярном и информационно насыщенном способе определения поставщика для государственных закупок. Для проверки эмпирической модели используются данные единой информационной системы (ЕИС) об электронных аукционах в 2014–2016 гг. и данные о решениях ФАС по делам о картельных сговорах.

## 1. МЕТОДЫ ВЫЯВЛЕНИЯ КАРТЕЛЕЙ

Сговоры на торгах представляют серьезный риск для эффективности государственных закупок. *Борьба со сговором на торгах* (bid-rigging) стала одним из важнейших направлений для антимонопольных органов во всем мире; этой проблеме были посвящены рекомендация ОЭСР 2012 г. и соответствующий доклад 2016 г. (OECD, 2016). Совершенствование закупочных процедур, их цифровизация помогают предотвращать и выявлять картели, но в то же время приводят к появлению новых антиконкурентных схем поведения. В данном разделе рассматриваются существующие методы выявления картелей, признаки сговора и условия его существования.

### 1.1. Поведенческий и структурный подходы

Методы обнаружения картелей делятся на структурные и поведенческие. Структурные предполагают выявление рынков с характеристиками, которые способствуют сговору, и проведение сфокусированного исследования работы фирм на этих рынках. Например, в работах (Symeonidis, 2003; Grout, Sonderegger, 2007) показано, что создание картеля вероятнее в отрасли с меньшим числом фирм, более однородной продукцией

и стабильным спросом. Во внимание принимаются три категории факторов:

- структурные факторы: число конкурентов, прозрачность рынка, барьеры входа;
- факторы предложения: степень однородности продукции, издержки фирм, отдача от инноваций;
- факторы спроса: колебания, эластичность и рост спроса, сильная покупательная способность.

Поведенческие методы направлены на выявление картелей путем анализа действий фирм на рынках и их координации между собой. Поведенческий подход также может быть сосредоточен на влиянии такой координации на рынок – на структуру цен, число фирм, динамику инвестиций и т.д. На подозрение о сговоре может навести параллельное движение цен или необоснованный их рост (Harrington, 2008; Levenstein, Suslow et al., 2004).

Указанные методы используются либо в *ex ante*<sup>1</sup>, либо в *ex post*<sup>2</sup> анализе картелей.

Обычно выявление картелей базируется на ответах на один из следующих вопросов (Harrington, 2008):

- Является ли поведение несовместимым с конкуренцией?
- Существует ли структурный разрыв в поведении фирм?
- Отличается ли поведение фирм, подозреваемых в сговоре, от поведения фирм-конкурентов?

*Поведение, характерное для конкуренции.* Ответ на первый вопрос базируется на выявлении свойств поведения, которые всегда имеются в условиях конкуренции или, по крайней мере, характерны для широкого класса конкурентных моделей, и проверке гипотезы их на-

<sup>1</sup> Анализ *ex ante* означает, что исследование проводится без предварительной информации о сговоре (Imhof et al., 2018; Chotibhongs, Arditi, 2012; Jakobsson, 2007).

<sup>2</sup> При анализе *ex post* информация о сговоре доступна, и исследователи могут разделить сделки на конкурентные и картельные (Porter, Zona, 1993, 1999; Pesendorfer, 2000).

личия для конкретной отрасли. Такой подход к тестированию имеется, например, в исследованиях Портера, Зоны (Porter, Zona, 1993, 1999) и Баджари и Йе (Bajari, Ye, 2003).

*Структурный разрыв в поведении.* Вторым методом выявления сговора – поиск структурного разрыва в поведении фирмы. Проверяется, изменились ли после предполагаемого момента создания (распада) картеля средняя цена, взаимосвязь между ценами фирм и т.д. (Connor, 2001; Abrantes-Metz, Froeb et al., 2006; Abrantes-Metz, Kraten et al., 2012). Так, в статьях Абрантес-Метц и соавторов и Болотовой и др. показано, что средняя цена на продукцию во время сговора возрастает, а стандартное отклонение цен, наоборот, уменьшается (Abrantes-Metz, Froeb et al., 2006; Bolotova, Connor et al., 2008).

Структурные изменения могут быть не связаны со сговором. Например, создание ассоциации может привести к усилению корреляции цен компаний, поскольку способствует обмену информацией, но неясно, приведет ли это к росту средних цен.

*Сравнение подозреваемых в сговоре с контрольной группой.* При сравнительном анализе в качестве бенчмарка могут выступать сопоставимые рынки, на которых фирмы не считаются сговорившимися, в частности различные географические рынки для данного продукта или услуги. Для панельных данных в качестве бенчмарка может выступать период до создания или после распада картеля. В одной статье российских авторов (Morozov, Podkolzina, 2013) выборка разделялась на основе разницы между ценой победителя и резервной ценой. Для каждой подвыборки строилась регрессия относительной цены контракта на число участников, число позиций (лотов) в закупке, объем текущих контрактов победителя и опыт победителя.

*Контроль над издержками фирмы.* При объяснении величины ставок участников торгов в качестве контрольных переменных чаще всего выделяют расстояние до места заклю-

чения контрактов, число выполняемых фирмами контрактов и специализацию каждой фирмы (Porter, Zona, 1993, 1999; Pesendorfer, 2000; Aryal, Gabrielli, 2013). Чем ближе фирма располагается к месту выполнения контракта, чем больше у нее свободных мощностей и чем точнее соответствует специализация фирмы объекту закупки, тем с большей вероятностью фирма выставит наименьшую ставку и выигрывает контракт.

Структурный и поведенческие методы дополняют друг друга, снижая тем самым вероятности ошибок первого и второго рода. Структурный метод часто дает слишком много ложных положительных результатов, например, для отраслей с высокими барьерами входа. Однако структурные показатели («экраны» – screens) дают мало ложных отрицательных результатов, поэтому они могут помочь исключить менее рискованные с точки зрения картелей отрасли и оставить только часть отраслей для использования поведенческих «экранов».

## 1.2. Методы машинного обучения

В последние годы для выявления картелей стали чаще применяться методы машинного обучения (Huber, Imhof, 2019; Ivanov, Nesterov, 2019).

Имхоф и Хубер (Huber, Imhof, 2019) используют простые статистические показатели для того, чтобы определить, как сговор влияет на распределение ставок. Они выдвигают несколько предположений и доказывают их на основе данных о крупном картеле в Тичино (Швейцария), существовавшем на протяжении нескольких лет. В частности, авторы показывают, что при сговоре на аукционе:

- снижается коэффициент вариации ставок<sup>3</sup>;
- распределение ставок становится более асимметричным из-за большей разни-

цы между первой и второй самыми низкими ставками.

Всего было рассчитано шесть статистических показателей, к которым в качестве контрольных были добавлены число участников и цена контракта. На основе этих показателей авторы применяют методы машинного обучения для создания моделей вероятности наличия картеля в закупке: лассо для логит-моделей и ансамблевый классификатор. В качестве показателей эффективности прогнозирования были взяты две меры: средние квадратичные ошибки прогнозируемых вероятностей сговора и доля правильных классификаций в тестовой выборке.

Машинное обучение также вошло в основу подхода, представленного в работе (Ivanov, Nesterov, 2019), где анализируются данные о более чем 1,4 млн запросов котировок на российском рынке в период с 2014 по 2018 гг. Авторы пытались выявить «утечку заявок» (bid leakage) – коррупционную схему, при которой поставщик незаконно предоставляет предпочитаемому участнику информацию о других заявках (Andreyanov et al., 2016). Результаты исследования показывают, что утечка заявок более вероятна на аукционах с более высокой резервной ценой, меньшим числом участников и меньшим снижением цен, а также в тех случаях, когда выигравшая заявка поступает в последний час до истечения крайнего срока.

В исследовании Иванова и Нестерова вводится важная предпосылка: проигравшие участники считаются честными, а победители – с какой-то вероятностью коррумпированными. Описываемая в статье процедура выявления утечки заявок состоит из двух этапов. На первом этапе обучается бинарный классификатор, чтобы отличать победителей от занявших второе место. С использованием кроссвалидации получают прогнозы этого классификатора для всех победителей в наборе данных. На втором этапе эти прогнозы преобразуются в вероятности утечки заявок.

<sup>3</sup> Отношение среднеквадратического отклонения к среднему значению ставки в аукционе.

## 2. СОСТОЯНИЕ КОНКУРЕНЦИИ НА ЭЛЕКТРОННЫХ АУКЦИОНАХ В РОССИИ И ПРАКТИКА АНТИМОНОПОЛЬНОЙ СЛУЖБЫ ВЫЯВЛЕНИЯ КАРТЕЛЕЙ

Государственные закупки на российском рынке осуществляются разными способами, которые регламентируются Федеральным законом № 44-ФЗ (Федеральный закон «О контрактной системе в сфере закупок товаров, работ, услуг для обеспечения государственных и муниципальных нужд» от 5 апреля 2013 г. № 44-ФЗ): 1) закупка у единственного поставщика; 2) запрос котировок; 3) запрос предложений; 4) электронный аукцион; 5) конкурс.

Особенности разных видов закупок обуславливают возможности применения разных схем и стратегий поведения, как вертикальных, так и горизонтальных (Balsevich, Podkolzina, 2014; Andreyanov et al., 2016). Не все схемы применимы ко всем видам закупок.

### 2.1. Общие тенденции в применении электронных аукционов как вида закупочных процедур

В нашем исследовании мы концентрируем внимание на открытых электронных аукционах (ОАЭФ) (ст. 59 № 44-ФЗ), которые с каждым годом становятся все более популярным способом определения поставщика для государственных закупок. Их доля в общем объеме закупок выросла с 55% в 2015 г. до 69% в 2018 г., а по числу извещений – с 57 до 60%, достигнув 1,96 млн извещений на общую сумму 5,66 трлн р.<sup>4</sup>

<sup>4</sup> Доклад о результатах мониторинга применения Федерального закона от 5 апреля 2013 г. № 44-ФЗ «О контрактной системе в сфере закупок товаров, работ, услуг для обеспечения государственных и муниципальных нужд» в 2015 г. URL: <http://economy.gov.ru/minec/about/structure/depfks.old/44/20160217>; Сводный аналитический отчет по результатам осу-

Эта динамика во многом объясняется рядом преимуществ электронного аукциона:

- информация о проводимых электронных аукционах хранится в открытом доступе, и производителям легко отслеживать предложения по закупкам;
- дистанционное взаимодействие с заказчиком без посредников уменьшает денежные и временные издержки;
- заявки анонимные, фирма-победитель называется только при оглашении результатов, что снижает коррупционные риски. Все этапы проведения электронного аукциона фиксируются, информация размещается на сайте электронной площадки, это увеличивает прозрачность торгов.

Тем не менее возможности для нечестных сделок сохраняются и в рамках электронного аукциона, и встречаются как вертикальные стоворы между заказчиком и поставщиком, так и горизонтальные между поставщиками. Электронные аукционы дают доступ к большему числу крупных государственных заказов, что усиливает стимулы участников торгов заключать антиконкурентные соглашения. Для того чтобы понять, какие стратегии поведения возможны со стороны оппортунистически настроенных фирм, рассмотрим элементы электронного аукциона.

### 2.2. Процедура проведения электронного аукциона в России

Процедура проведения ОАЭФ регламентируется Федеральным законом №44-ФЗ.

1. В единой информационной системе (ЕИС) в открытом доступе размещается извещение о проведении закупки и документация.

2. На сбор заявок для аукционов с начальной максимальной ценой контракта (НМЦК) меньше 3 млн р. участникам дается

\_\_\_\_\_ ществления мониторинга закупок, товаров, услуг для обеспечения государственных и муниципальных нужд в соответствии с Федеральным законом от 5 апреля 2013 г. № 44-ФЗ... по итогам 2018 г.

минимум 7 дней, для аукционов с большей НМЦК – минимум 15 дней.

3. Заказчик рассматривает первые части заявок и готовит протокол рассмотрения заявок (Протокол 1), где указано, какие поставщики допускаются до аукциона, а какие – нет.

4. Аукцион проходит в первый рабочий день после двух дней с момента завершения рассмотрения первых частей заявок. На подачу ценовых предложений отводится 10 мин, после каждого предложения время продлевается на 10 мин. Если в течение 10 мин после подачи последнего ценового предложения никто из участников больше не сделал более низкую ставку, то аукцион завершается. Участники делают ценовые предложения, снижая текущую цену в пределах «шага аукциона», который составляет 0,5–5% от НМЦК (ст. 68 № 44-ФЗ). После завершения аукциона в течение 30 минут публикуется протокол о проведении аукциона (Протокол 2).

5. После проведения аукциона в течение трех рабочих дней рассматриваются вторые части заявок, после чего принимается решение о соответствии/несоответствии фирм требованиям аукциона и публикуется протокол рассмотрения вторых частей заявок (Протокол 3). Только на этом этапе называются фирмы-участники.

6. После рассмотрения вторых частей заявок подводятся итоги аукциона, определяется победитель и подписывается контракт.

### 2.3. Антиконтурные действия на электронных аукционах

Форма электронного аукциона позволяет участникам сговоров на торгах применять несколько различных стратегий.

При горизонтальном сговоре, как правило, фирма-победитель выбирается картелем заранее, а на аукционе участники картеля делают фиктивные ставки, чтобы создать иллюзию конкуренции. Похожая ситуация на торгах происходит, если создается фиктивная компания, которая участвует в аукционе, но

никогда реально не конкурирует (поддельные торги). Схемы при горизонтальном сговоре.

1. *Поддержание цены на максимально высоком уровне.* Участники картеля делают ценовые предложения с минимально возможным понижением цены, и на этом аукцион завершается; некоторые участники могут не делать ставок. Цена, как правило, снижается всего на 0,5–1%. Такая схема применима и удобна, когда кроме участников картеля на торгах больше нет фирм. Для того чтобы аукцион состоялся, достаточно всего двух участников в картеле и аукционе. Если в отрасли довольно высокая конкуренция, большое число новых участников рынка, то отсутствие конкурентов на торгах может быть обеспечено путем дополнительного вертикального сговора с заказчиком. В последнее время стали применяться «аукционные роботы», которых программируют на минимальное снижение цены<sup>5</sup>.

2. *«Таран».* Еще одной популярной стратегией является «таран», при котором добросовестные участники вытесняются с торгов<sup>6</sup>. Два или несколько членов картеля активно торгуются, демпингуют и этим поведением отпугивают добросовестных участников, которые понимают, что не могут так сильно снижать цену, и в какой-то момент останавливаются. Еще один член картеля делает ставку чуть ниже добросовестных участников, но выше «таранов». По итогам рассмотрения вторых частей заявок оказывается, что фирмы, которые играли роль «таранов», не проходят по требованиям к документации,

<sup>5</sup> ФАС в СМИ: ФАС нашла признаки картеля на торгах с использованием аукционных роботов. URL: <https://fas.gov.ru/publications/15211>; ФАС в СМИ: ФАС доказала неправомерное использование аукционных роботов при закупках расходных материалов. URL: <https://fas.gov.ru/publications/15258>.

<sup>6</sup> Доклад «Способы осуществления государственных закупок и административная ответственность за нарушение законодательства о контрактной системе». URL: <https://fas.gov.ru/documents/612473>.

и в итоге выигрывает участник картеля, сделавший более высокую ставку.

Основные схемы антиконкурентных действий заказчика или организации, проводящей закупку, при вертикальном сговоре:

- сдерживание подачи заявок (написание названия закупки с ошибками, усложнение контрактной документации, чрезмерно жесткие требования к участникам, слишком сжатые сроки выполнения контракта);
- отказ в приеме заявки;
- отмена закупки.

## 2.4. Практика выявления антимонопольной службой сговоров на торгах

Большинство дел о картелях, возбуждаемых ФАС России, относятся к сговорам на торгах (в 2018 г. – 86% дел о картелях). Чаще всего сговоры выявляются в ремонте и строительстве, торговле лекарственными средствами, продуктами питания, на транспорте и в сделках с недвижимостью<sup>7</sup>.

На основании опыта разбирательств, инициированных ФАС, можно выделить следующие доказательства наличия сговора между хозяйствующими субъектами:

- минимальное снижение НМЦК;
- одинаковые:
  - IP-адреса, с которых фирмы отправляют заявки;
  - электронные адреса, почты, телефоны фирм;
  - учетные записи, под которыми созданы заявки;
  - размеры файлов;
  - дата/время создания, внесения изменений заявок;
- сходство текстов заявок, аффилированность участников:
  - идентичность лицензионных документов;
  - идентичность торговых наименований;

- общие учредители;
- совпадение или близкое расположение фактических адресов компаний;
- связи со сторонней организацией, которая, возможно, координирует участие в торгах;
- договоренности о получении квалифицированного сертификата ключа проверки электронной подписи;
- заключение договоров субподряда между участниками;
- устойчивые финансовые связи между участниками.

## 3. ДАННЫЕ

В данном исследовании мы представляем метод обнаружения картелей, в основе которого лежит ex-post анализ сговоров на закупках. Иными словами, модель строится с учетом данных о том, на каких аукционах был обнаружен сговор. В качестве источников данных для модели были использованы Единая информационная система государственных закупок (ЕИСГС) и сайт ФАС.

Стратегия выявления картелей состоит из двух шагов. Сначала на основе данных о закупках и о решениях ФАС определяются характерные признаки аукционов, на которых были выявлены картели. Затем с помощью машинного обучения тренируется классификатор, который для каждого аукциона определяет вероятность участия в нем картеля.

Сбор и обработка данных состояли из нескольких этапов:

- 1) получение исходной выборки;
- 2) предварительная обработка;
- 3) построение матрицы документов и фильтрация по шаблону;
- 4) извлечение данных из объединенного XML-документа.

Информация о закупках, в которых участвовал картель, была получена из решений по делам ФАС об антиконкурентных соглашениях за период с 2014 по 2016 г. Извлечение

<sup>7</sup> Доклад о состоянии конкуренции в РФ за 2018 г. URL: <https://fas.gov.ru/documents/685117>.

номеров закупок проводилось по регулярным выражениям. В исходной выборке 214 дел, по которым признавались антиконкурентные соглашения на 5055 закупочных процедурах. Для контрольной группы закупок было случайным образом отобрано 20 тыс. аукционов из всего массива закупок в 2014–2016 гг. На следующих этапах были отброшены аукционы, если по ним не было части необходимых документов (извещений и протоколов), если эти документы были пустыми или содержали ошибки.<sup>8</sup>

В итоге осталось 5101 наблюдение из случайной выборки и 2789 наблюдений из выборки ФАС. В табл. 1 приводится описательная статистика по закупкам, где обнаружен картель (выборка ФАС) и где не обнаружен (случайная выборка).

#### 4. СТРАТЕГИЯ ОЦЕНИВАНИЯ И ГИПОТЕЗЫ

В данной статье для выявления картельных сговоров мы применяем статистические показатели процесса закупки в сочетании с машинным обучением. Для проведения различия между сговором и конкурентным поведением участников аукциона рассматривается несколько переменных, построенных на основе распределения заявок и ставок в каждой закупке. Для каждой заявки выдвигается гипотеза о связи ее размера с наличием картеля. В частности, делаются предположения относительно ставок участников, времени подачи заявок, числа участников и их активности на торгах, длительности этапов аукциона, а также максимальной начальной цены аукциона.

<sup>8</sup> Подробное описание сбора и обработки данных содержится в приложении 1. Все приложения представлены в качестве дополнительного материала в электронной версии статьи на сайте журнала <https://www.ecr-journal.ru>

Ниже описаны показатели, использованные в анализе.

*Разница во времени подачи заявок.* Как показывает практика ФАС, заявки от участников сговора часто подаются с одного IP-адреса, а документы к заявкам создаются с использованием одной и той же учетной записи. Это может означать, что одно и то же лицо (или группа лиц) занимается подготовкой и отправкой заявок нескольких компаний, и поэтому документы фирм – участников картеля с большей вероятностью будут отправлены в одно и то же время. Даже если фирмы готовят и отправляют документы самостоятельно, согласованность их действий может повлиять на то, что разница во времени подачи заявок будет небольшой. Для того чтобы учесть статистическую связь между временем подачи заявки и наличием картеля, было посчитано несколько показателей.

Первая переменная –  $mean\_dift\_app_j$  – средняя разница во времени подачи заявки между участниками  $i$  и  $(i - 1)$  аукциона (участник  $(i - 1)$  – фирма, которая подала заявку последней перед участником  $i$ ).

*Гипотеза 1.* Вероятность картеля выше в аукционе, в котором меньше средняя разница во времени подачи заявки между участниками  $i$  и  $(i - 1)$ .

Вторая переменная –  $median\_dift\_app_j$  – медиана разницы во времени подачи заявки между участником  $i$  и  $(i - 1)$  аукциона (участник  $(i - 1)$  – фирма, которая подала заявку последней перед участником  $i$ ).

*Гипотеза 2.* Вероятность картеля выше в аукционе, где меньше медиана разницы во времени подачи заявки между участниками  $i$  и  $(i - 1)$ .

В выборке с картелями распределение разницы во времени между заявками участников больше смещено к нулю, чем в случайной выборке (рис. 1). Для картелей доля закупок, в которых  $median\_dift\_app_j$  лежит в пределах одного часа, в 2 раза больше, чем в случайной выборке. Это подтверждает выдвинутые гипотезы.

Таблица 1

Описательная статистика процедур проведения электронных аукционов

Переменная	Mean	Std. deviation	Min	Max
NO CARTELS ( $N = 5101$ )				
Число заявок по Протоколу 1	3,02	1,49	2,00	17,00
Число заявок по Протоколу 2	2,60	1,04	2,00	14,00
Разница в числе заявок между Протоколами 1 и 2	0,42	1,10	0,00	15,00
День недели публикации извещения	1,99	1,50	0,00	6,00
Медиана разницы между временем подачи заявки и началом приема заявок, ч	24,21	42,52	0,00	660,23
Медиана разницы во времени подачи заявки между $i$ и $(i - 1)$ участниками, ч	239,03	208,12	2,12	5508,02
НМЦК, тыс. р.	5643,33	40910,57	0,07	1616001,00
Длительность торгов, мин	16,60	30,36	0,00	406,42
Максимальное число ставок	2,93	6,69	1,00	80,00
Разница между минимальной ценой и НМЦК, % НМЦК	15	18	1	80
Разница между минимальной и второй ценами, % НМЦК	0,7	0,8	0	5
Среднее снижение цены за один шаг, % НМЦК	4,5	6	0,1	39
Число опубликованных документов	5,68	1,5	5	27
CARTELS ( $N = 2789$ )				
Число заявок по Протоколу 1	2,33	0,86	2,00	11,00
Число заявок по Протоколу 2	2,09	0,38	2,00	7,00
Разница в числе заявок между Протоколами 1 и 2	0,25	0,75	0,00	7,00
День недели публикации извещения	2,05	1,49	0,00	6,00
Медиана разницы во времени подачи заявки и началом приема заявок, ч	9,80	26,19	0,00	293,02
Медиана разницы во времени подачи заявки между $i$ и $(i - 1)$ участниками, ч	233,80	131,74	2,61	1463,90
НМЦК, тыс. р.	6481,77	27812,92	1,06	734362,60
Длительность торгов, мин	4,63	12,04	0,00	149,15
Максимальное число ставок	1,18	2,13	1,00	49,00
Разница между минимальной ценой и НМЦК, % НМЦК	4	10	1	74
Разница между минимальной и второй ценами, % НМЦК	0,5	0,4	0	7,2
Среднее снижение цены за один шаг, % НМЦК	1	2	0,2	27
Число опубликованных документов	5,32	0,88	5	19

Источники: рассчитано авторами.

Разница между временем подачи заявки участником  $i$  и началом сбора заявок, скорее всего, не будет связана с вероятностью победы участника  $i$  и наличием картеля. В электронном аукционе ставки делаются в отдельный день, поэтому не важно, в какой момент подана заявка. Из следующего рисун-

ка (рис. 2) видно, что значительных отличий между двумя выборками по этому показателю нет, как и предполагалось. Два пика в области 7-го (144–168 часов) и 15-го дня (336–360 часов) свидетельствуют о том, что многие участники подают заявки ближе к концу срока их сбора.

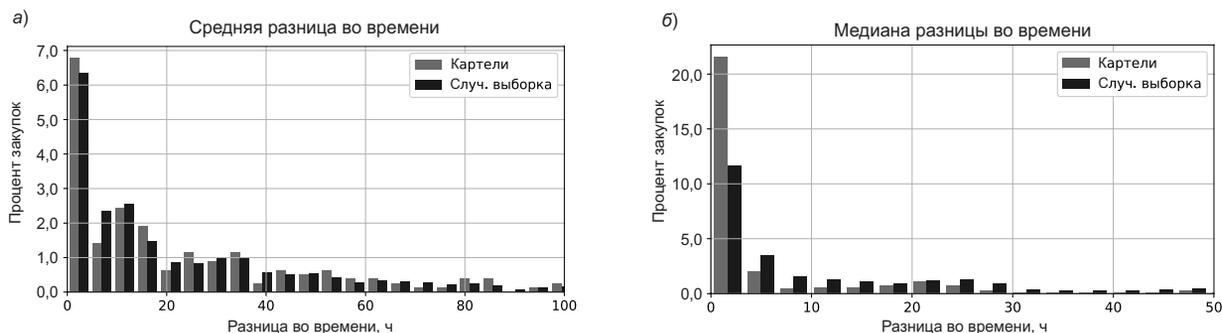


Рис. 1. Распределение средней разницы (а) и медианы (б) во времени подачи заявок

Источники: рассчитано авторами.

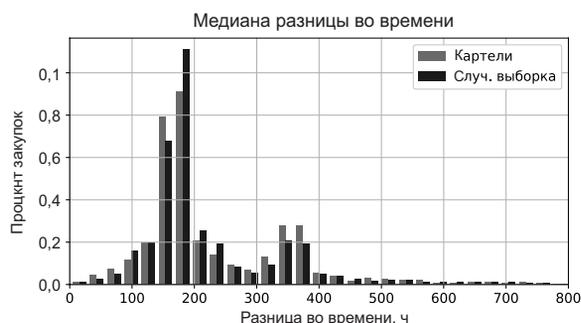


Рис. 2. Распределение разницы во времени между подачей заявки и началом сбора заявок

Источники: рассчитано авторами.

Если в торгах участвует картель, то с большой вероятностью выиграет один из его участников. При этом, как отмечалось выше, если фирмы находятся в сговоре, то с большой вероятностью их заявки будут отправлены на электронную торговую площадку почти одновременно. Поэтому на вероятность наличия картеля в аукционе  $j$  может влиять такая фиктивная переменная  $win\_and\_mint_j$ , что если

$$\min(t_{app_i} - t_{app_{i-1}}) = t_{app_k} - t_{app_{k-1}}, \text{ то}$$

$$win\_and\_mint_k =$$

$$= \begin{cases} 1, & \text{если участник } k \text{ или} \\ & (k-1) \text{ стал победителем;} \\ 0, & \text{если участник } k \text{ или} \\ & (k-1) \text{ не стал победителем.} \end{cases}$$

*Гипотеза 3.* Если в закупке стал победителем один из участников, чья заявка имела минимальную разницу во времени с предыдущей или последующей, то вероятность картеля в таком аукционе больше.

По данным выборки (для аукционов с числом участников больше двух) получилось, что вероятность ( $win\_and\_mint_j = 1$ ) в картелях – 85%, а в случайной выборке – 68%.

*Длительность сбора заявок.* В случае сговора между заказчиком и поставщиками, его участникам может быть выгодно провести аукцион в максимально сжатые сроки. Чем короче срок сбора заявок, тем ниже вероятность того, что нежелательные конкуренты увидят извещение о торгах, составят и подадут заявку. Графики распределения для наших данных не показали значительных отличий в длительности сбора заявок между выборками по торгам, где выявлен картель, и где не выявлен. Это можно объяснить тем, что честным заказчикам тоже не выгодно затягивать проведение аукциона. Связи между наличием картеля и днем недели, в который было опубликовано извещение, тоже не было найдено<sup>9</sup>.

*Активность участников на аукционе.* Если аукцион проходит в условиях честной конкуренции

<sup>9</sup> Связи между наличием картеля и днем недели, в который было опубликовано извещение, тоже не было найдено (см. приложение 2).

ции, то участники будут торговаться активнее, это отразится на числе ставок, времени проведения аукциона и размере снижения цены.

После начала аукциона участникам дается 10 мин на то, чтобы сделать ставки, и после каждого ценового предложения время автоматически продлевается. Поэтому, если фирмы активно конкурируют за контракт, продолжительность аукциона увеличивается. В качестве показателя берем разницу во времени между первым ценовым предложением и последним, т.е. минимальной ставкой:

$$Dt_{\text{auction}} = t_{\text{winner}}(\text{last\_date}) - t_{\text{firstbid}}(\text{first\_date}).$$

*Гипотеза 4.* Чем дольше длился аукцион, тем ниже вероятность наличия в нем картеля.

График распределения длительности аукциона (рис. 3) подтверждает наше предположение. Всплеск доли аукционов в выборке по картелям на отметке в 10 мин говорит о том, что часто последняя ставка делалась непосредственно перед окончанием срока, отведенного на новое ценовое предложение.

Следующий показатель активности торгов – максимальное число ставок, которое сделал участник в аукционе  $j$ :

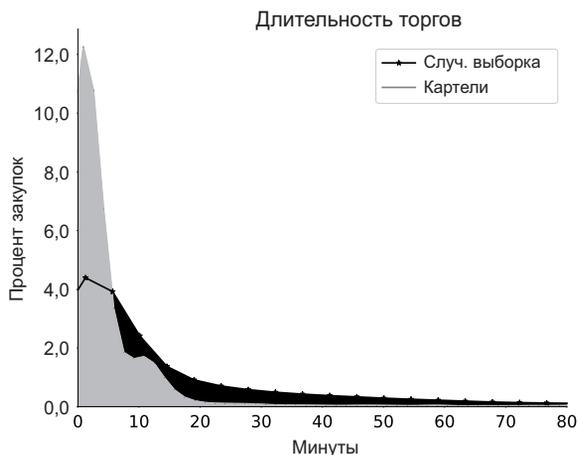


Рис. 3. График распределения длительности проведения электронного аукциона

Источники: рассчитано авторами.

$$\max\_countbids = \max_i(\text{offers\_quantity}_i).$$

*Гипотеза 5.* Чем больше максимальное число ставок, которое сделал участник в аукционе, тем ниже вероятность наличия картеля.

Участники делали по одной ставке в 97,3% закупок, где был выявлен картель, и в 84,3% случайной выборки. Нет почти ни одного аукциона с картелем, где хотя бы один участник сделал больше пяти ставок, для случайной выборки – ситуация иная (рис. 4). Такие результаты логичны. Если на торгах есть сговор, то победитель, как правило, заранее определен и нужно делать ставки, только чтобы аукцион не был отменен.

*Разница в ценах.* При наличии картельного сговора цена в ходе торгов снижается намного меньше при прочих равных условиях. Для того чтобы обеспечить сопоставимость закупок, при анализе используем относительное снижение цены, взвешиваем разницу по начальной максимальной цене контракта (НМЦК):

$$MaxMinP_j = \frac{\max\_price_j - \min\_price_{ij}}{\max\_price_j},$$

где  $\max\_price_j$  – НМЦК на аукционе  $j$ ,  $\min\_price_{ij}$  – минимальная цена на аукционе  $j$ .

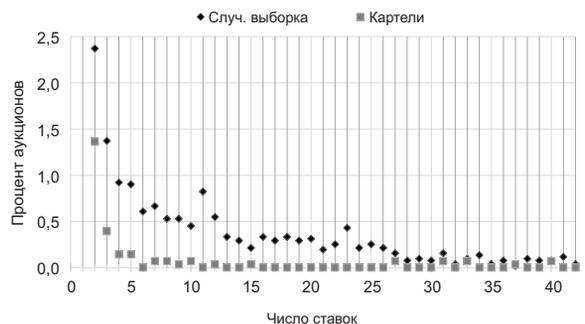


Рис. 4. Плотность распределения максимального числа ставок, сделанных участником

Источники: рассчитано авторами.

*Гипотеза 6.* Чем меньше разница между минимальной ценой и НМЦК, тем выше вероятность наличия картеля.

В выборке аукционов с картелями распределение разницы между минимальной и резервной ценами очень сжато, в большинстве наблюдений разница составляет менее 5% НМЦК (рис. 5).

Если вторая самая низкая ставка близка к минимальной цене, то закупочное агентство может заключить контракт с фирмой, сделавшей второе наилучшее ценовое предложение, если другие критерии компенсируют разницу в цене. Участники картеля могут предпочесть сохранить определенную разницу между первой и второй самыми низкими ставками, чтобы гарантировать определенный исход, желательный для картеля:

$$SecondMinP_j = \frac{secondbest\_price_j - win\_price_j}{max\_price_j}$$

*Гипотеза 7.* Чем больше разница между минимальной и второй по величине ценой, тем больше вероятность наличия картеля.

Есть всего несколько значений, которые принимает переменная  $SecondMinP_j$

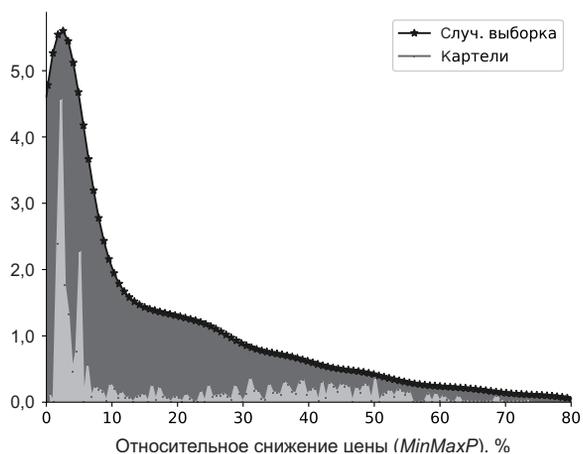


Рис. 5. Плотность распределения разницы между максимальной и минимальной ценами, % от максимальной цены

Источники: рассчитано авторами.

Таблица 2

Распределение разницы между минимальной и второй по величине ценой, %

Значение $SecondMinP_j$	Доля закупок в случайной выборке	Доля закупок в выборке по картелям
< 0,5	7,92	11,87
0,5	76,81	83,04
(0,5; 5)	12,92	4,88
5	2,35	0,18
> 5	0,00	0,04

Источники: рассчитано авторами.

почти во всех аукционах, доля остальных значений близка к нулю, поэтому удобнее будет показать отличия между выборками на основе следующей таблицы (табл. 2), а не графика распределения. В обеих выборках  $SecondMinP_j$  в большинстве аукционов равняется 0,5%, т.е. нижней границе «шага аукциона», установленного № 44-ФЗ. В случайной выборке оказывается больше доля тех аукционов, где  $SecondMinP_j \in (0,5\%; 5\%)$  или равно 5%. Наличие значений меньше «шага аукциона» можно объяснить ч. 12 ст. 68 № 44-ФЗ. Что касается аукционов, где  $SecondMinP_j > 5\%$ , то таких наблюдений оказалось меньше 0,1% объема выборки<sup>10</sup>.

Конечно, на цены влияет масса различных факторов, важно учитывать объект закупки, макроэкономические условия, государственную политику в отношении отраслей и т.д. Кроме того, разные отрасли характеризуются разным уровнем конкуренции в це-

<sup>10</sup> В выборке есть аукционы, для которых не соблюдался «шаг аукциона», установленный № 44-ФЗ (ст. 68). Эта ситуация требует дополнительного исследования. Ее можно объяснять либо ошибками в данных, либо какими-то дополнительными условиями закупки. В основной документации закупок, которой мы располагаем, не содержится информации по поводу «шага аукциона». Для итоговой выборки не отделяются аукционы, в которых не соблюдается интервал шага (0,5; 5), так как это сильно уменьшает выборку и снижает точность классификаторов.

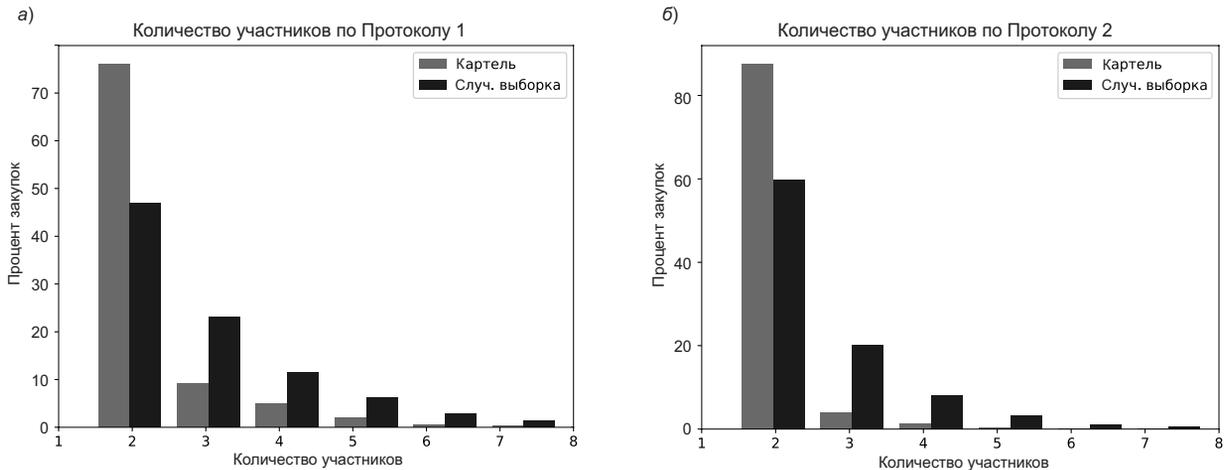


Рис. 6. Плотность распределения числа участников аукциона по Протоколам 1 (а) и 2 (б)

Источники: рассчитано авторами.

лом. Число участников и снижение цены также могут зависеть от числа позиций (лотов) в закупке. Чем больше товарных позиций, тем меньше компаний готовы выполнить его из-за ограничений специализации или мощностей фирмы.

**Число участников.** Чем большее число участников в аукционе, тем труднее осуществить стратегию картеля по поддержанию цены. С ростом числа участников, скорее всего, будет расти активность конкуренции, и больше снижаться цена в ходе торгов.

**Гипотеза 8.** С ростом числа участников аукциона снижается вероятность наличия картеля.

Для учета числа участников были использованы показатели:

$count\_app\_pr1_j$  – число подавших заявку на аукцион  $j$ ;

$count\_app\_pr2_j$  – число допущенных до электронного аукциона  $j$ .

В 77% выборки по картелям заявки на участие в аукционе подавали всего две фирмы. В случайной выборке аукционов доля закупок, состоящих из двух участников, – около 46% (рис. 6, а). Картелю выгодно, когда на торгах нет конкурентов-аутсайдеров. Если на

рынке они есть, то картель может сдержать их появление на аукционе путем сговора с заказчиком. Если же сторонние компании все-таки, решили участвовать, заказчик может не допустить их к торгам по первой части заявок. Участник картеля может способствовать отклонению заявок конкурентов<sup>11</sup>. По нашим данным (рис. 6, б) в случае наличия картеля доля аукционов с двумя участниками – около 90%, в случае их отсутствия – 60%.

Значительные различия можно увидеть также в совместном распределении показателей в двух выборках. Например, в случайной выборке наблюдается положительная связь между числом ставок, сделанных одним участником, и относительным снижением цены. Это означает, что участники аукциона активно торгуются, вследствие чего цена сильно уменьшается – причем как для крупных, так и для небольших контрактов (рис. 7). В выборке по картелям почти во всех аукционах участники делают только одно ценовое

<sup>11</sup> Пример такого поведения содержится в деле ФАС № 1-11-123/00-22-17: поставщик, находившийся в сговоре с заказчиком, отправлял заказчику письма, где указывал основания, по которым нужно отклонить некоторые заявки.

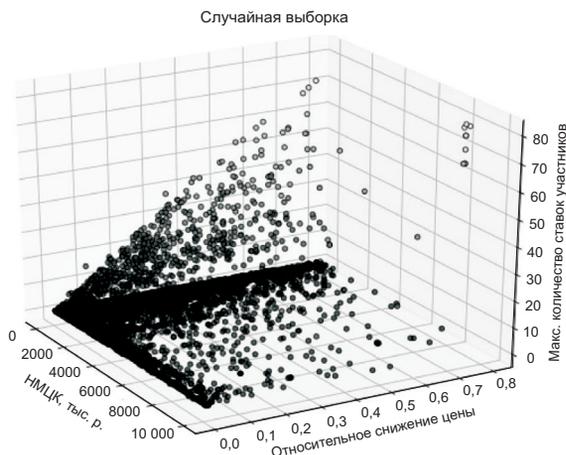


Рис. 7. Совместное распределение НМЦК, снижения цены в ходе аукциона и максимального числа ставок. Случайная выборка

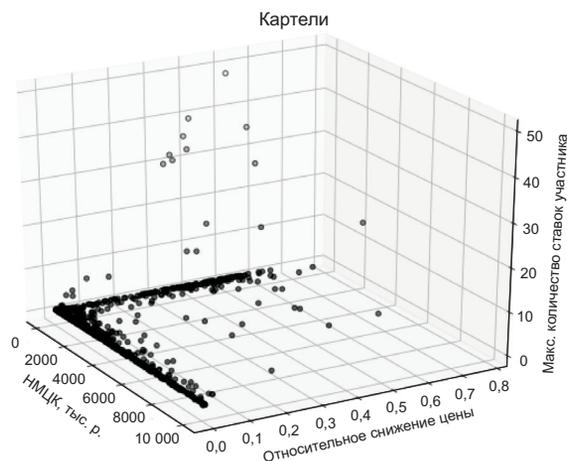


Рис. 8. Совместное распределение НМЦК, снижения цены в ходе аукциона и максимального числа ставок. Выборка по картелям. Методы машинного обучения

предложение (рис. 8), а цена значительно снижается, только когда НМЦК контракта небольшая, меньше 200 тыс. р.<sup>12</sup>

## 5. МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ

Мы применяем методы машинного обучения для тренировки и тестирования модели, которая предсказывает вероятность существования картеля на электронном аукционе. Используется несколько различных моделей, в том числе два простых классификатора – логистическая регрессия и метод ближайших соседей и два ансамблевых – случайный лес и градиентный бустинг (Gradient Boosting).

Каждый классификатор распределяет аукционы на «честные» и «нечестные», основываясь на данных о восьми переменных:

- $count\_app\_pr2_i$  – число участников, допущенных к аукциону  $i$ ;
- $dif\_countapp12_i$  – разница между числом фирм, подавших заявки, и числом фирм, допущенных к аукциону  $i$ ;

- $median\_dif_i$  – медиана разницы во времени подачи заявок попарно взятых участников, идущих друг за другом по времени на аукционе  $i$ ;

- $maxP\_thous_i$  – максимальная начальная цена контракта (НМЦК) на аукционе  $i$ , тыс. р.,

- $xD\_min\_maxPrice_i$  – разница между НМЦК и минимальной ставкой на аукционе  $i$ ;

- $wD\_Second\_Min\_Price_i$  – разница между второй наилучшей ставкой и минимальной ставкой на аукционе  $i$ ;

- $max\_countbids_i$  – максимальное число ценовых предложений, сделанных участником на аукционе  $i$ ;

- $count\_doc\_clarification_i$  – число опубликованных поясняющих документов для аукциона  $i$ . Предполагаем, что большое число разъяснений может говорить о том, что участвовали игроки, которые не были в сговоре, а пытались уточнить требования, чтобы оценить свои шансы на победу.

Выбор модели, которая предсказывает вероятность существования картеля на электронном аукционе, осуществлялся в несколько этапов.

Сначала были сформированы обучающая и тестовая выборки. Наша первоначаль-

<sup>12</sup> Более подробная информация о взаимосвязи переменных содержится в приложении 3.

ная выборка не была сбалансирована: в ней много наблюдений, где не выявлен картель, и этот класс модель будет определять точнее. Для того чтобы сгладить несбалансированность, используем SMOTE (Synthetic Minority Oversampling Technique) алгоритм (Chawla et al., 2002), который строит синтетические наблюдения для миноритарного класса, т.е. для закупок с картелями. После применения SMOTE каждый класс стал занимать 50% выборки, а общее число наблюдений увеличилось до 10202.

Затем 70% наблюдений было отнесено к обучающей выборке, которая использовалась для оценки параметров модели. Тестовая выборка из 30% наблюдений применялась для прогнозирования и оценки эффективности. Далее данные были стандартизированы, чтобы снизить влияние масштаба данных на оценки классификаторов.

Для каждой модели были подобраны оптимальные показатели<sup>13</sup>, и затем классификаторы сравнивались между собой по критериям качества. Качество моделей оценивалось с помощью:

- правильности предсказания (Accuracy)<sup>14</sup>;
- точности классификации (Precision)<sup>15</sup>;
- полноты классификации (Recall)<sup>16</sup>.

<sup>13</sup> Некоторые расчеты оптимальных параметров содержатся в приложении 4.

<sup>14</sup> Правильность предсказания (Accuracy) показывает долю правильно предсказанных классификаций, т.е. разделение на закупки с картелями и без них.

<sup>15</sup> Точность (Precision) показывает, какая доля из тех наблюдений, которые помечены как положительные (т.е. принимается гипотеза о принадлежности к определенному классу), действительно принадлежат этому классу.

<sup>16</sup> Полнота (Recall) показывает, какую долю наблюдений, относящихся к определенному классу, классификатор отнес к этому классу.

## 6. РЕЗУЛЬТАТЫ

Правильность предсказания каждой из моделей, построенной с 10 разбиениями кроссвалидации всей выборки, равняется:

- Accuracy: 0,74 (+/-0,02) [Логистическая регрессия];
- Accuracy: 0,84 (+/-0,05) [Случайный лес];
- Accuracy: 0,78 (+/-0,02) [KNN соседи];
- Accuracy: 0,84 (+/-0,04) [Градиентный бустинг],

где в скобках указана стандартная ошибка для показателя Accuracy.

Результаты показывают, что логистическая модель хуже остальных определяет наличие или отсутствие картеля. Наиболее эффективными являются случайный лес и градиентный бустинг. Такие же выводы получаются и при анализе ROC-кривых, которые отображают соотношение доли наблюдений, для которых правильно показано отсутствие картеля (True positive rate), и наблюдений, для которых неверно показано отсутствие картеля (False positive rate) (рис. 9).

Для оценки эффективности моделей, определяющих классы cartel («есть картель»)

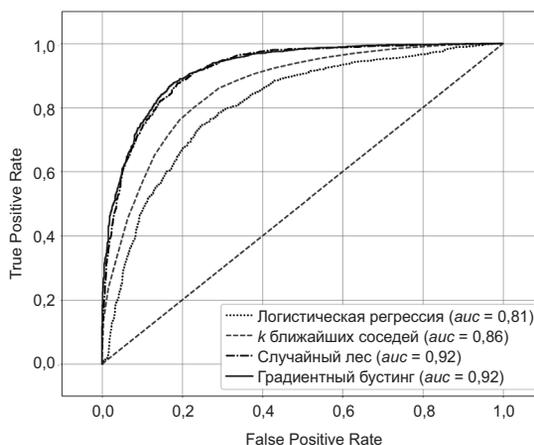


Рис. 9. ROC-кривые для моделей выявления картелей

Источники: рассчитано авторами.

Таблица 3

Точность и полнота моделей, показывающих наличие или отсутствие картеля на электронном аукционе

Модель	Наличие картеля	Precision	Recall	f1-score	N (тестовая выборка)
Логистическая модель	Нет картеля	0,83	0,58	0,68	1523
	Картель	0,68	0,88	0,76	1538
<i>k</i> ближайших соседей	Нет картеля	0,80	0,77	0,78	1523
	Картель	0,78	0,81	0,79	1538
Случайный лес (100 деревьев)	Нет картеля	0,86	0,81	0,84	1523
	Картель	0,83	0,87	0,85	1538
Градиентный бустинг	Нет картеля	0,88	0,80	0,84	1523
	Картель	0,82	0,89	0,85	1538

Источник: рассчитано авторами.

и `no_cartel` («нет картеля»), используем показатели точности и полноты модели (табл. 3). Ансамблевые модели (случайный лес и градиентный бустинг) наиболее эффективны, обе модели строятся на основе деревьев, но отличаются порядком и способом объединения результатов. Для случайного леса 87% (Recall) всех закупок из тестовой выборки, где был обнаружен картель, были правильно отнесены классификатором к неконкурентным. И 83% (Precision) закупок, отнесенных классификатором к закупкам с картелем, действительно были признаны таковыми ФАС.

Для случайного леса на основе матрицы ошибок классификации были посчитаны вероятности ошибок первого и второго рода<sup>17</sup> в тестовой выборке: их процент – 6,01 и 9,11% соответственно. Также была рассчитана относительная важность признаков (табл. 4). Наибольший вклад в определение наличия картеля на аукционе вносят уровень снижения цены, разница во времени подачи заявок и величина НМЦК.

<sup>17</sup> Для нашей задачи ошибки первого рода – наблюдения, в которых неверно предсказано наличие картеля, ошибки второго рода – наблюдения, в которых неверно предсказано отсутствие картеля.

## ЗАКЛЮЧЕНИЕ

В данной работе разрабатывался подход к выявлению горизонтальных сговоров на государственных закупках, осуществляемых через открытые электронные аукционы. Эмпирические модели строились на основе данных о закупках с сайта ЕИС и данных по картелям из решений дел ФАС. С помощью методов машинного обучения были построены классификаторы, которые предсказывают наличие или отсутствие картеля для каждой закупки в зависимости от распределения ценовых предложений участников аукциона,

Таблица 4

Относительная важность признаков

Переменная	Относительная важность признака
<i>w_D_min_maxPrice</i>	0,33
<i>meadian_dift_hours</i>	0,21
<i>maxP_thous</i>	0,16
<i>w_D_Second_Min_Price</i>	0,12
<i>count_app_pr2</i>	0,08
<i>count_doc_clarification</i>	0,04
<i>dif_countapp12</i>	0,035
<i>max_countbids</i>	0,02

Источник: рассчитано авторами.

времени подачи заявок, длительности проведения торгов и числа участников.

Результаты исследования показали, что модели случайного леса и градиентного бустинга наиболее точно предсказывают выявление картеля на электронном аукционе. Правильность предсказания в обеих моделях равняется 84%, а полнота и точность – 0,83 и 0,87 в случайном лесе и 0,82 и 0,89 – в градиентном бустинге.

Важно отметить, что выводы о наличии картеля на закупке очень сильно зависят от качества работы ФАС и особенностей их деятельности, так как мы не знаем истинное число картелей и ориентируемся только на стговоры, выявленные антимонопольной службой.

С помощью классификаторов можно посчитать для каждого аукциона вероятность обнаружения в нем картеля. Эта информация может быть использована не только контролирующими органами, но и нарушителями. Участники картелей могут скорректировать свою стратегию таким образом, чтобы их поведение на аукционе стало менее подозрительным. Тогда пороги отнесения аукциона к честному/нечестному снова придется менять. Поиск равновесия и выявление оптимальных стратегий контролирующего органа и нарушителя являются предметом нашей следующей статьи, где проблема участия картеля в государственных закупках формулируется в рамках теории игр как игра с инспекцией.

Представленный подход к выявлению стговоров может быть усовершенствован в дальнейших исследованиях. Данные о предмете закупок, разделение аукционов по отраслям позволят контролировать общий уровень конкуренции в отраслях, характерные черты закупок для разных отраслей (объем и прибыльность контрактов, снижение цены в ходе торгов, случайные шоки спроса и предложения для разных товаров и т.д.). Это улучшит качество моделей. Данные о предмете закупок могут быть получены в первую очередь с помощью анализа методами машинной лингвистики текстов извещений об электронном аукционе и заявок участников. Еще улучшить

предсказательную силу модели можно благодаря информации об аффилированности участников аукциона. Прокси для показателя аффилированности могут быть построены на основе анализа текстов заявок, а также данных о связях фирм, которые можно получить из баз корпоративной информации.

## Список литературы / References

- Abrantes-Metz R.M., Froeb L.M., Geweke J.F., Taylor C.T.* A variance screen for collusion // *International Journal of Industrial Organization*. 2006. No. 24. P. 467–486.
- Abrantes-Metz R.M., Kraten M., Metz A.D., Seow G.* Labor manipulation // *Journal of Banking and Finance*. 2012. No. 36. P. 136–150.
- Andreyanov P., Davidson A., Korovkin V.* Corruption vs Collusion: Evidence from Russian Procurement Auctions // *Technical Report*. 2016. mimeo: UCLA.
- Aryal G., Gabrielli M.F.* Testing for collusion in asymmetric first-price auctions // *International Journal of Industrial Organization*. 2013. Vol. 31. P. 26–35.
- Bajari P., Ye L.* Deciding between competition and collusion // *The Review of Economics and Statistics*. 2003. No. 85. P. 971–989.
- Balsevich A., Podkolzina E.A.* Indicators of corruption in public procurement: The example of Russian regions // *Higher School of Economics*. 2014. Research Paper No. WP BRP 76/EC/2014. URL: Available at SSRN: <https://ssrn.com/abstract=2530518> or <http://dx.doi.org/10.2139/ssrn.2530518>.
- Bolotova Y., Connor J., Miller D.* The Impact of collusion on price behavior: Empirical results from two recent cases // *International Journal of Industrial Organization*. 2008. No. 26. P. 1290–1307.
- Chawla N.V., Bowyer K.W., Hall L.O., Kegelmeyer W.P.* SMOTE: Synthetic minority over-sampling technique // *Journal of Artificial Intelligence Research*. 2002. Vol. 16. P. 321–357.
- Chotibhongs R., Arditi D.* Analysis of collusive bidding behavior // *Construction Management and Economics*. 2012. No. 30. P. 221–231.

- Connor J.M. Global price fixing: Our customers are the enemy. Boston: Kluwer Academic Publishers, 2001.
- Grout P., Sonderegger S. Structural approaches to cartel detection. In: E. C-D, Atanasiu I. (eds.). European Competition Law Annual: 2006. Enforcement of Prohibition of Cartels. Hart Publishing, 2007. P. 83–103.
- Harrington J.E. Detecting cartels. In: Handbook of Antitrust Economics, Buccirosi P. (ed.). Cambridge: The MIT Press, 2008.
- Huber M., Imhof D. Machine learning with screens for detecting bid-rigging cartels // International Journal of Industrial Organization. 2019. Vol. 65. P. 277–301.
- Imhof D., Karagoek Y., Rutz S. Screening for bid rigging, does it work? // Journal of Competition Law and Economics. 2018. No. 14 (2). P. 235–261.
- Ivanov D.I., Nesterov A.S. Identifying bid leakage in procurement auctions: Machine learning approach. 2019. arXiv:1903.00261 [econGN].
- Jakobsson M. Bid rigging in Swedish procurement auctions. Working Paper. Mimeo: Uppsala University. Department of Economics (Sweden), 2007.
- Levenstein M.C., Suslow V.Y., Oswald L.J. Contemporary international cartels and developing countries: Economic effects and implications for competition policy // Antitrust Law Journal. 2004. No. 71. P. 801–852.
- Morozov I., Podkolzina E. Collusion detection in procurement auctions // SSRN Electronic Journal. 2013. 10.2139/ssrn.2221809.
- OECD, 2016. Fighting bid rigging in public procurement: Report on implementing the OECD Recommendation. URL: <http://www.oecd.org/daf/competition/Fighting-bid-rigging-in-public-procurement-report-2016.pdf>.
- Pesendorfer M. A study of collusion in first-price auction // The Review of Economic Studies. 2000. No. 67. P. 381–411.
- Porter R.H., Zona J.D. Detection of bid rigging in procurement auctions // The Journal of Political Economy. 1993. No. 101. P. 518–538.
- Porter R.H., Zona J.D. Ohio school milk markets: An analysis of bidding // RAND Journal of Economics. 1999. No. 30. P. 263–288.

Symeonidis G. In which industries is collusion more likely? Evidence from the UK // Journal of Industrial Economics. 2003. No. 51. P. 45–74.

Рукопись поступила в редакцию 16.08.2019 г.

## DETECTING INDICATORS OF HORIZONTAL COLLUSION IN PUBLIC PROCUREMENT WITH MACHINE LEARNING METHODS

G.O. Molchanova, A.I. Rey, D.Yu. Shagarov

DOI: 10.33293/1609-1442-2020-1(88)-109-127

*Glafira O. Molchanova*, Institute for Industrial economics RANEPa, Moscow, Russia; ORCID 0000-0001-8130-1259; molchanova-go@ranepa.ru

*Alexey I. Rey*, Institute for Industrial economics RANEPa, Moscow, Russia; ORCID 0000-0001-8207-1790; rey-ai@ranepa.ru

*Dmitriy Yu. Shagarov*, Institute for Industrial economics RANEPa, Moscow, Russia; ORCID 0000-0001-6544-5359; shagarov-dy@ranepa.ru

Improvement of procurement procedures and their digitization help prevent and identify cartels, but at the same time lead to the emergence of new anticompetitive schemes. In this paper we focus on electronic auctions, which have become the main method of public procurement in Russia in recent years. As e-auctions provide access to many big government orders; the incentives for bidders to join anticompetitive agreements are increased. Therefore, the development of methods to detect bid rigging at electronic auctions is of high practical importance. The aim of this work was to develop a method for detecting signs of horizontal collusion at an auction. We use machine learning methods to train classifiers that predict the presence or absence of cartel in electronic auctions, depending on the distribution of bidders, the time of submission of applications, the duration of the auction and the number of participants. Variables for the model were selected on the basis of distribution plots

---

built for sample of cartels and random sample. The study is based on data from public procurement Web portal and the information about bid rigging from cases of the Federal Antimonopoly Service. The results showed that the Random forest model most accurately predicts the detection of the cartels on electronic auctions. The accuracy of the prediction is 84%, and the recall and precision of the model are 83 and 87%, respectively. The most significant variables for the classification are the level of price reduction, the difference in the time of application filing of participants and the value of the maximum starting price of contract.

*Keywords:* detecting cartels, bid rigging, procurement auctions, screening methods, machine learning.

*JEL:* C12, C38, C63, D22, D44, H57, L40, L41.

*Manuscript received 16.08.2019*

---

## ОБ ОСОБЕННОСТЯХ ВОСПРИЯТИЯ ПРЕДПРИНИМАТЕЛЬСТВА СТУДЕНТАМИ ВУЗОВ

*Н.А. Винокурова, Н.М. Светлов*

**DOI:** 10.33293/1609-1442-2020-1(88)-127-142

На материалах опросов студенческой молодежи в двух университетах г. Воронежа (680 респондентов) отклонена гипотеза о том, что студенты в согласии с трактовкой предпринимательской деятельности в трудах Й. Шумпетера связывают предпринимательство с инновационной деятельностью. Установлено, что восприятие предпринимательства студентами в российской культурно-лингвистической среде отличается от шумпетерианского. Интерес студенчества к тому, чтобы в будущем заняться предпринимательством, не свидетельствует о его готовности к новаторству. Высокий предпринимательский потенциал учащейся молодежи сочетается с ее низкой готовностью к инновационному предпринимательству.

Выявлено, что у студенческой молодежи отсутствует целостная система взглядов и приоритетов в отношении карьеры. Преобладают личные интересы и мотивации, выражающиеся в их предпочтениях относительно места будущей работы, где главным условием названа возможность сочетать работу и личную жизнь, в их стремлении к финансовой безопасности, к материальному благополучию, к уклонению от трудностей. Работа предпочтается лишь интересная и творческая. Деловые наклонности слабы, они уступают личностным интересам. Трудолюбие как фактор успеха в будущей профессио-

---

© Винокурова Н.А., Светлов Н.М., 2020 г.

*Винокурова Наталья Анатольевна*, старший научный сотрудник, Центральный экономико-математический институт РАН, Москва, Россия; [vinokurova@yandex.ru](mailto:vinokurova@yandex.ru)  
*Светлов Николай Михайлович*, ведущий научный сотрудник, Центральный экономико-математический институт РАН, Москва, Россия; ORCID 0000-0001-6906-6129; [nikolai.svetlov@gmail.com](mailto:nikolai.svetlov@gmail.com)