ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА И ГОСУДАРСТВЕННОЙ
СЛУЖБЫ ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ» (РАНХИГС)

ОЦЕНКА И ПРОГНОЗИРОВАНИЕ ПОСЛЕДСТВИЙ ВЫХОДА РОССИЙСКИХ ФИРМ НА ЭКСПОРТНЫЕ РЫНКИ С УЧЕТОМ ЭФФЕКТА ОТБОРА

Москва 2021

Ключевые слова: ПРОДУКТИВНОСТЬ ФИРМЫ, МЕЖДУНАРОДНАЯ ТОРГОВЛЯ, МОДЕЛЬ СТОХАСТИЧЕСКОЙ ГРАНИЦЫ, МОДЕЛЬ ВЫСОКОЙ РАЗМЕРНОСТИ, ОЦЕНКА ЭФФЕКТА ВОЗДЕЙСТВИЯ

ВВЕДЕНИЕ

Необходимым условием для обеспечения устойчивого экономического роста российской экономики является увеличение эффективности предприятий. Одним из факторов, который может к этому привести, является расширение вовлечённости России в международную торговлю. Наиболее очевидными являются последствия эффектов, связанных с ростом числа более эффективных компаний-экспортеров в результате возросшей международной специализации. Предприятия с положительным экспортным статусом, как правило, более крупные, поэтому при увеличении уровня вовлечённости страны в международную торговлю производственные ресурсы перетекают в сторону более производительных компаний экспортёров, что приводит к увеличению общего уровня эффективности экономики. В то же время, в рамках гипотезы об обучении посредством экспорта, выход новых компаний на экспортные рынки будет вести к увеличению их эффективности, а значит и росту эффективности российской экономики в целом.

Итак, рост экономики можно связать с тремя закономерностями: во-первых, увеличивается число экспортирующих фирм, во-вторых, при начале экспортной деятельности, согласно предположению об обучении через экспортную деятельность, фирмы, которые вышли на внешний рынок, становятся более эффективными, в-третьих, предприятия-экспортёры укрупняются в размере, что приводит к перетеканию факторов производства к более эффективным фирмам с положительным экспортным статусом.

В данной работе исследуется гипотеза об обучении посредством экспорта. Задача усложняется тем, что прямое измерение приводит к системным ошибкам. Выход на внешние рынки связан с преодолением барьеров, что приводит к дополнительным издержкам для предприятий. Получить выгоду от такого рода деятельности могут только изначально более производительные фирмы, что соответствует эффекту отбора. Получается, что при простом сравнении фирм с различным экспортным статусом, экспортёры окажутся более эффективными, но при этом невозможно будет сказать какова доля эффекта обучения посредством экспорта в данном различии.

1 Закономерности и последствия отбора компаний на экспортные рынки

1.2 Эмпирические исследования эффекта отбора

В статье [1], авторы рассматривали разницу в уровне производительности между фирмами с различными экспортными статусами, на основе данных обрабатывающих предприятий Испании с 1991 по 1996 годы. Также в ней рассмотрены две гипотезы, описывающие отличие фирм-экспортёров от фирм, работающих только на внутреннем рынке: гипотеза об отборе более производительных фирм на внешние рынки и гипотеза об обучении фирм при их взаимодействии с зарубежными партнёрами.

Существует предположение о наличии границы производственных возможностей, которая показывает максимальный уровень производства при заданных факторах на данном этапе развития технологий. Немногие исследователи ставят задачу о её точном определении. Едва ли такая задача вообще является разрешимой. Однако само её понятие позволяет говорить об удалённости фирмы от границы производственных возможностей. Если две аналогичные фирмы с близким размером факторов производства выпускают отличающиеся объёмы продукции, то, очевидно, та фирма, выпуск которой выше, находится ближе к границе своих производственных возможностей.

Среди фирм существует существенное различие в их уровне производительности [2]. Существенной характеристикой, влияющей на различие в уровне эффективности фирм, является их экспортный статус: фирмы-экспортёры находятся ближе к своей границе производственных возможностей, чем фирмы, оперирующие только на внутреннем рынке [3], [4], [5], [6], [7].

Выделяют два объяснения различий в уровне продуктивности между фирмами с различным экспортным статусом: во-первых, фирмы, работающие на внешних рынках, сталкиваются с более сильной конкуренцией, что вынуждает их оптимизировать своё производство, во-вторых, они несут невозвратные издержки от выхода на экспорт, что отсеивает наименее продуктивные фирмы. Оба объяснения говорят о том, что при выходе на экспорт происходит селекция наиболее производительных фирм. Фирмы сталкиваются с более серьёзной конкуренцией на внешних рынках, соответственно у наименее производительных компаний меньше шансов продолжить своё существование на данных рынках. Из-за появления невозвратных издержек от выхода на экспорт наименее эффективные фирмы не могут себе позволить совершить необходимые для выхода на экспортные рынки инвестиции. В данном случае, даже если предположить, что уровень конкуренции на внутреннем и на внешних рынках совпадают, присутствие невозвратных издержек всё равно приведёт к различию в уровне эффективности между фирмами с различным экспортным статусом.

Если рассматривать не фиксированные во времени данные, где можно сравнивать фирмы с различным экспортным статусом, но ещё и обратиться к динамике их изменения, то можно строить модель, описывающую динамику фирм. Модель, предполагающая наличие невозвратных издержек от выхода на экспортные рынки, приводит к тому, что экспортный статус каждой компании в фиксированный момент времени, сильно связан с её экспортным статусом в предыдущий момент.

Присутствие невозвратных издержек от выхода на экспортные рынки приводит к тому, что фирмы, которые решают выйти на мировой рынок, вынуждены нести дополнительные потери, относительно фирм, которые остаются работать только на внешнем рынке. Таким образом, только изначально более эффективные предприятия будут выходить на внешние рынки. Соответственно изначально все экспортёры являются более производительными, чем фирмы, оставшиеся на внутреннем рынке. Также, после выхода на внешние рынки фирмы сталкиваются с усиленной конкуренцией, что приводит к тому, что наименее эффективные экспортёры будут вынуждены уйти с внешних рынков.

Предположения о присутствии невозвратных издержек и увеличенном уровне конкуренции на внешних рынках отражают эффект отсеивания наименее производительных фирм от участия в международной торговле.

Кроме того, существует предположение, что фирмы, вышедшие на экспортные рынки, повышают свою эффективность вследствие взаимодействия с иностранными партнёрами. Происходит восприятие других технологий, новых практик производства, обмен опытом работы, и другое. Данные эффекты называют обучением через экспорт.

Оценить данные эффекты можно разными способами. В работе [1], рассматриваются четыре процедуры, которые должны были выявить описанные выше эффекты. Первая заключается в том, что, если распределение фирм-экспортёров по уровню производительности должно быть выше, чем у фирм, оперирующих только на внутреннем рынке.

Следующая процедура рассматривает сравнение фирм, которые меняли свой экспортный статус с отрицательного на положительный, с фирмами, которые сохранили свой отрицательный экспортный статус. Если фирмы меняли свой экспортный статус с отрицательного на положительный в момент времени $t + 1$, то сравнение производится в предшествующий момент времени t. Распределение по производительности фирм, менявших свой экспортный статус, должно доминировать распределение фирм, которые сохранили отрицательный экспортный статус.

Третья процедура представляет собой обратную к первой. В ней сравнивались фирмы, которые, меняли свой экспортный статус с положительного на отрицательный, с теми фирмами, которые сохранили свой экспортный статус. Таким же образом, как и в первой процедуре, в ней сравниваются распределения фирм в момент времени t, где смена статуса происходила в момент времени $t+1$. В данном случае, распределение по уровню эффективности фирм, которые сохранили свой экспортный статус, должно доминировать распределение фирм, которые в следующий момент времени поменяли свой экспортный статус с положительного на отрицательный.

Последняя, четвёртая процедура, использовалась для того, чтобы оценить эффект обучения посредством экспортной деятельности. Чтобы его уловить, отслеживались фирмы, которые сменили свой экспортный статус в определённый момент времени. Если данное предположение верно, то после начала экспорта, их отрыв по уровню производительности, от фирм, которые не начали экспортировать, должен увеличиться. Если данный эффект не присутствует, то фирмы, после выхода на внешние рынки, продолжают работать также, как они работали до выхода на экспорт, и, таким образом, их различие по уровню производительности от фирм, оставшихся на внутреннем рынке, не должно существенно меняться.

Панельные данные позволяют отслеживать траекторию разных фирм во времени, и сравнивать их различные характеристики, что позволяет применять процедуры, описанные выше. Фирмы можно разбить на необходимые группы, на основе истории их экспортного статуса. После измерения их продуктивности, можно делать вывод о влиянии экспортного статуса на производительность фирм. Большинство описанных процедур сводится к сравнению различных групп фирм, на основе их экспортного статуса. Сравнение групп основано на понятии стохастического доминирования распределений, которое определяет порядок на множестве различных групп фирм. Если распределение уровня производительности фирм в одной группе равно F, а в другой – G, то первая группа превосходит вторую по распределению, если $F(z) - G(z) \leq 0$, где z отвечает за уровень производительности фирм.

Введём некоторые обозначения, чтобы можно было более строго сформулировать описанные выше процедуры оценивания влияния экспортного статуса на уровень производительности. Пусть $Z_1, \dots, Z_n$ обозначают определённую характеристику фирм в первой группе (например уровень производительности фирм или рост уровня производительности), функция распределения которой равна F. А значения данной характеристики для фирм в другой группе с функцией распределения G обозначим за $Z_{n+1}, \dots, Z_{n+m}$. Тогда интересующие гипотезы можно сформулировать в виде тестов.

В первом тесте основную гипотезу можно сформулировать в виде:

$$H_0: F(z) - G(z) = 0, \quad \forall z \in \mathbb{R} \tag{1}$$

где H_0 – обозначение основной гипотезы,

F – функция распределения первой группы,

G – функция распределения второй группы,

z – наблюдаемый параметр (уровень производительности или рост уровня производительности).

Альтернативная гипотеза будет записана в виде:

$$H_1: F(z_0) - G(z_0) \neq 0, z_0 \in \mathbb{R} \tag{2}$$

где H_1 – обозначение альтернативной гипотезы,

F – функция распределения первой группы,

G – функция распределения второй группы,

z_0 – фиксированное значение наблюдаемого параметра (уровень производительности или рост уровня производительности).

Формулы (1) и (2) описывают нулевую и альтернативную гипотезы в тесте. В первом из них значение наблюдаемого параметра может принимать любое действительное значение. По сути, гипотеза состоит в том, что распределения F и G совпадают. Утверждение второй гипотезы заключается в том, что распределения F и G различны, для этого необходимо, чтоб они различались хотя бы в одной какой-либо точке z_0.

Другой тест можно сформулировать в виде дух гипотез. Нулевая гипотеза имеет вид:

$$H_0: F(z) - G(z) \leq 0, \quad z \in \mathbb{R}, \tag{3}$$

где H_0 – обозначение основной гипотезы,

F – функция распределения первой группы,

G – функция распределения второй группы,

z – наблюдаемый параметр (уровень производительности или рост уровня производительности).

Альтернативная гипотеза запишется в виде:

$$H_1: F(z_0) - G(z_0) > 0, \quad z_0 \in \mathbb{R}, \tag{4}$$

где H_1 – обозначение альтернативной гипотезы,

F – функция распределения первой группы,

G – функция распределения второй группы,

z_0 – фиксированное значение наблюдаемого параметра (уровень производительности или рост уровня производительности).

В уравнениях (3), (4) проверяется гипотеза о доминировании распределения G над распределением F. Основная гипотеза заключается в том, что при всех значениях наблюдаемого параметра z, распределение $F(z)$ не больше распределения $G(z)$. Альтернативная гипотеза говорит о том, что существует такое значение параметра $z = z_0$, при котором значение функции распределения $F(z_0)$ больше значения функции распределения $G(z_0)$.

Описанные выше гипотезы можно сформулировать, используя супремум. Тогда двусторонний тест можно записать в виде следующих гипотез. Нулевая гипотеза:

$$H_0: \sup_{z \in \mathbb{R}} |F(z) - G(z)| = 0, \tag{5}$$

где H_0 – обозначение основной гипотезы,

F – функция распределения первой группы,

G – функция распределения второй группы,

z – наблюдаемый параметр (уровень производительности или рост уровня производительности).

Альтернативная гипотеза примет вид:

$$H_1: \sup_{z \in \mathbb{R}} |F(z) - G(z)| \neq 0, \tag{6}$$

где H_1 – обозначение альтернативной гипотезы,

F – функция распределения первой группы,

G – функция распределения второй группы,

z_0 – фиксированное значение наблюдаемого параметра (уровень производительности или рост уровня производительности).

Формулы (5) и (6) формулируют туже гипотезу, что и формулы (1) и (2) только в другом виде. Данный тест также проверяет, равны ли между собой распределения двух групп, где функция распределения первой группы равна F, а функция распределения второй группы равна G. Соответственно, нулевая гипотеза говорит о том, что супремум модуля разности данных функций на всех возможных значениях наблюдаемого параметра z равен нулю. Альтернативная гипотеза утверждает, что супремум модуля разности функций распределения F и G на всех возможных значениях наблюдаемого параметра z не равно нулю. Это соответсвует ситуации, когда существует такое значение $z = z_0$, при котором функции $F(z)$ и $G(z)$ не совпадают.

Теперь перейдём к другой формулировке одностороннего теста. Его нулевая гипотеза может быть сформулирована в виде:

$$H_0: \sup_{z \in \mathbb{R}} (F(z) - G(z)) \leq 0, \tag{7}$$

где H_0 – обозначение основной гипотезы,

F – функция распределения первой группы,

G – функция распределения второй группы,

z – наблюдаемый параметр (уровень производительности или рост уровня производительности).

Альтернативная гипотеза примет вид:

$$H_1: \sup_{z \in \mathbb{R}} (F(z) - G(z)) > 0, \tag{8}$$

где H_1 – обозначение альтернативной гипотезы,

F – функция распределения первой группы,

G – функция распределения второй группы,

z_0 – фиксированное значение наблюдаемого параметра (уровень производительности или рост уровня производительности).

Формулы (7) и (8) представляют в другом виде гипотезы, сформулированные в формулах (3), (4). Данный тест проверяет, доминирует ли распределение, заданное функцией распределения $G(z)$, распределение, заданное функцией распределения $F(z)$. Если верна нулевая гипотеза, то при всех возможных значениях наблюдаемого параметра z, значения функции распределения $F(z)$ не больше значений функции $G(z)$. Если же верна альтернативная гипотеза, то существует хотябы одна точка $z = z_0$, где значение функции распределения $F(z_0)$ превышает значение функции распределения $G(z_0)$.

Данные, по которым проверялись приведённые выше гипотезы в работе [1], состояли из опроса испанских фирм. Данный опрос подходил по-разному к фирмам разного размера. В данном опросе отделили крупные фирмы, они определялись как фирмы, в которых нанято больше двухсот сотрудников. Крупные фирмы просили участвовать в опросе. Согласилось около семидесяти процентов от общего числа всех крупных фирм страны. Для малых фирм, в которых было нанято от десяти до двухсот сотрудников, опрос проводился иначе: опрашиваемые фирмы выбирались случайно. Общий размер выборки составил около пяти процентов всех малых фирм. Таким образом, данные покрывают разные доли фирм, в зависимости от их размера. Соответственно, исследователям было необходимо контролировать на размер компании.

Другой особенностью доступных исследователям данных было то, что они не представляли собой сбалансированную панель. Это связано с тем, что от года к году фирмы, которые участвовали в опросе, в нём сохранялись, а новые, ранее не опрошенные фирмы, добавлялись. Поэтому количество измерений в кросс-секции каждого года менялось от одного года к другому.

Данные собирались в период с 1991 года по 1996. Всего было собрано 10595 наблюдений, что соответствует в среднем 1766 наблюдениям в каждый год. Все фирмы в каждый год разбивались на пять групп. В одну группу входили фирмы, у которых на протяжении всего исследуемого периода сохранялся положительный экспортный статус. Во вторую группу поместили фирмы, у которых на протяжении всего исследуемого периода сохранялся

отрицательный экспортный статус. В третью группу поместили фирмы, которые меняли свой экспортный статус с отрицательного на положительный в рассматриваемый период. В четвёртую группу поместили фирмы, которые меняли свой экспортный статус с положительного на отрицательный в рассматриваемый период. В пятую группу вошли фирмы, которые на рассматриваемом периоде неоднократно меняли свой экспортный статус.

Было обнаружено, что при прямом сравнении двух групп с разным экспортным статусом фирмы-экспортёры оказались более крупными. Среди крупных фирм в приведённых двух группах 78% фирм являются большими. Среди маленьких фирм данный показатель составил только 27%. Сравнение групп фирм, которые однократно меняли свой экспортный статус с положительного на отрицательный, или с отрицательного на положительный в течении данного периода показывает, что значительная доля фирм меняет свой экспортный статус. Доля малых фирм, которые меняли свой экспортный статус, составила 16%. Доля крупных фирм, которые меняли свой экспортный статус составила 11%. Причём, за исследуемый период доля фирм, которые меняли свой экспортный статус с отрицательного на положительный, превышает долю фирм, которые меняли свой экспортный статус с положительного на отрицательный. Таким образом, исследователи связывают рост экспорта Испании с 1991 по 1996 годы с ростом числа экспортирующих фирм. В последнюю группу фирм, которые меняли свой экспортный статус неоднократно, вошло 11% малых фирм и 6% крупных.

Эффективность фирмы не является прямо наблюдаемой величиной. Чтобы о ней говорить, нужно договориться, каким образом она определяется с помощью наблюдаемых величин. Тогда к данным можно добавить ещё одну переменную, отвечающую за уровень производительности компаний, и использовать её для оценки тестов, предложенных выше гипотез. В качестве производительности был использован индекс, который строится на основании данных о количестве выпускаемой фирмой продукции, количестве используемого труда, капитала и расходных материалов.

Выпуск фирмы определялся как суммарный годовой доход фирмы за её товары и услуги в реальном выражении. Труд определялся как суммарное количество часов, которые сотрудники работали в течении года. Используемые материалы измеряли в качестве расходов фирм на используемые услуги других фирм, закупку необходимого сырья, топлива и так далее, расходы на материалы представлялись в реальном выражении. Капитал измерялся с помощью формулы:

$$k_t^* = I_t + k_{t-1}^*(1 - d_t)\frac{P_t}{P_{t-1}}, \quad (9)$$

где t – индекс года измерения,

k_t^* – размер капитала фирмы в год t,

I_t – инвестиции в капитал в год t,

d_t – амортизация капитала в год t,

P_t – цены на оборудование в год t.

Формула (9) показывает, как меняется капитал фирмы от года к году. Она учитывает амортизацию оборудования, посредством коэффициента d_t. Соответственно, она учитывает изменение капитала вследствие его устаревания и инфляции, за которую отвечает последняя дробь. Так же капитал может изменяться в результате инвестиций, что учитывает слагаемое I.

Доступные данные можно представить в виде множества:

$$\left\{\left(Y_{ft}, W_{ft}^k, X_{ft}^k, k = 1, \dots, K\right), f = 1, \dots, N, t = 1, \dots, T\right\}, \quad (10)$$

где f – индекс фирмы,

t – индекс года измерения,

Y_{ft} – выпуск фирмы f в год t,

k – индекс материала,

W_{ft}^k – доля затрат от обещего дохода фирмы f на материал k в год t,

X_{ft}^k – количество затрачиваемого материала k фирмой f в год t,

K – число учитываемых материалов,

N – общее количество фирм в данных,

T – число доступных годов наблюдений.

Выражение (10) перечисляет доступные переменные, по которым строился индекс производительности фирм. Используя данные обозначения, можно построить индекс производительности в виде:

$$\ln\lambda_{ft} = \ln Y_{ft} - \overline{\ln Y}_{tr} - \frac{1}{2}\sum_{k=1}^{K}\left(W_{ft}^{k} + \overline{W}_{\tau\tau}^{k}\right)\left(\ln X_{ft}^{k} - \overline{\ln X}_{\tau r}^{k}\right) + \\ + \overline{\ln Y}_{\tau r} - \overline{\ln Y}_{r} - \frac{1}{2}\sum_{k=1}^{K}\left(\overline{W}_{\tau r}^{k} + \overline{W}_{r}^{k}\right)\left(\overline{\ln X}_{\tau r}^{k} - \overline{\ln X}_{r}^{k}\right), \quad (11)$$

где f – индекс фирмы в данных,

t – индекс года,

λ_{ft} – индекс производительности фирмы f в год t,

k – индекс материала,

τ – индикатор принадлежности фирмы к группе, где разбиение основанно на размере фирмы,

r – индикатор принадлежности фирмы к группе, где разбиение основанно на виде экономической деятельности,

W_{ft}^{k} – доля затрат от общего дохода фирмы f на материал k в год t,

X_{ft}^{k} – количество затрачиваемого материала k фирмой f в год t,

K – число учитываемых материалов.

Формула (11) показывает индекс производительности, по которому оценивался уровень эффективности фирм. Усреднение берётся по всем фирмам и всем годам наблюдений внутри каждой группы, описываемой индикаторами r и τ, в отдельности. Слагаемые в первой строке показывают на сколько фирма отличается от средней фирмы в своей группе, как по выпуску продукции, так и по затратам на данный выпуск. Второе слагаемое сравнивает различие фирмы в своей группе, основанное на размере от средней фирмы в экономике.

Для тестирования гипотез о статистическом доминировании экспортёров над не экспортёрами сравнивались фирмы из одинаковых групп, основанных на размере. В доступных исследователям данных выборка больших фирм существенно отличалась от выборки малых фирм.

Результаты оценки показали, что фирмы с положительным экспортным статусом действительно превосходят по распределению производительности фирмы с отрицательным экспортным статусом. Особенно ярко выражено данное различие у малых

фирм. Оценка эффекта обучения вследствие экспорта продукции у них получилась положительной, но в доверительный интервал также вошли значения, при которых данный эффект отсутствует. Поэтому однозначно говорить о данном эффекте, основываясь на используемой работе, нельзя.

Для исследования эффекта отбора используются полупараметрические методы оценки стохастической границы. Но для начала необходимы процедуры для оценки распределения наблюдаемых величин. Данные процедуры основаны на приближении функций с помощью некоторого базиса в функциональном пространстве. В данной работе в качестве базиса, с помощью которого приближаются функциональные зависимости, используются полиномы Лагерра:

$$\rho_m(w) = \sum_{l=0}^{m} C_m^l \frac{(-1)^l}{l!} w^l, \tag{12}$$

где w – переменная, от которой зависит приближаемая функция,

$\rho_m(w)$ – полином Лагерра степени m,

l – индекс, по которому идёт суммирование,

C_m^l – биномиальный коэффициент.

Численно каждая функция аппроксимируется с помощью вектора чисел, обозначим его за δ. С помощью формулы (12), можно вывеси соотношение для приближения плотности вероятности распределения. Расширив вектор δ добавлением единицы, таким образом, что $\xi = (1, \delta_1, \delta_2, \dots, \delta_n)'$, можно выписать выражения для оценки:

$$f(w|\delta) = \frac{\xi_{n+1}(\delta)' D_{n+1}(w) \xi_{n+1}(\delta)}{\xi_{n+1}(\delta)' \xi_{n+1}(\delta)}, \tag{13}$$

где w – переменная, от которой зависит приближаемая функция,

δ – вектор чисел, с помощью которого описывается приближение,

f – функция плотности распределения,

ξ – вспомогательный вектор, получающийся добавлением единицы к вектору δ,

D – вспомогательная матрица,

$n + 1$ – размерность векторов и матрицы, которая является квадратной.

В формуле (13) используется матрица D, выражение для неё можно записать в виде:

$$d_{k,m}(w) = \sum_{i=0}^{k}\sum_{j=0}^{m} C_k^i C_m^j \frac{(-1)^{i+j}}{i!\,j!} w^{i+j} exp(-w), \quad (14)$$

где w – переменная, от которой зависит приближаемая функция,

$d_{k,m}$ – элемент матрицы D, который стоит на k-ой строке и m-ом столбце,

i, j – индексы, по которым идёт суммирование,

C_k^i, C_m^j – биномиальные коэффициенты.

Как говорилось выше, приближение каждой функции задаётся вектором параметров δ. Соответственно, задача о приближении функции распределения сводится к задаче о нахождении коэффициентов δ. Используя формулы (13) и (14), была построена процедура для приближения функции плотности распределения. Ниже приведен ряд тестов, с помощью которых проверялась состоятельность данного приближения.

В таблице 1 представлены результаты аппроксимации экспоненциального распределения с параметром λ=2. Расстояние между функциями определялось как супремум разности модуля между функциями на заданном интервале, в численном случае сводилось к поиску максимума модуля разности среди точек, в которых оценивалась приближаемая функция. Как видно, данное простое распределение незначительно зависит от количества точек, в которых оценивается функция. Ошибка значительно уменьшается с ростом количества варьируемых параметров, с помощью которых задаётся оценочная функция: с увеличением количества параметров с двух до трёх ошибка уменьшается на порядок. Также видно, что ошибка возрастает с ростом количества точек, в которых проводится оценка функции, что соответствует ожиданию: для того, чтобы приблизить в меньшем количестве точек требуется меньшая точность. Как видно из результатов, данный метод аппроксимирует экспоненциальное распределение с большой точностью, уже при пяти варьируемых параметрах, ошибка составляет меньше одной тысячной.

Таблица 1 – Ошибка аппроксимации функции плотности экспоненциального распределения

количество точек	dim(**δ**)=1	dim(**δ**)=3	dim(**δ**)=5
11	0.046	0.003	0
101	0.052	0.004	0
1001	0.052	0.004	0

Примечание – функция аппроксимировалась с помощью полиномов Лагерра, приближение проводилось выбором элементов вектора **δ**. Числа в таблице показывают максимум модуля разности между фактическим значением функции плотности распределения и её аппроксимацией среди всех выбранных точек, в которых оценивалась функция плотности распределения. Первый столбец показывает количество точек, в которых оценивалась функция плотности распределения. Остальные столбцы показывают близость приближения в зависимости от количества варьируемых параметров, которые задавались с помощью вектора **δ**.

В таблице 2 приведены результаты приближения гамма распределения с параметром формы равным трём и параметром масштаба равным двум. Рассматривался отрезок от нуля до двадцати, с равномерным разбиением на десять, сто и тысячу отрезков, значение функции аппроксимировалось на концах этих отрезков. Как видно из таблицы, точность приближения существенно улучшается с увеличением количества варьируемых параметров. Интересно, что при увеличении числа точек измерения с одиннадцати до ста одной, ошибка уменьшается, хотя в большинстве случаев в таких задачах она должна иметь обратную зависимость. Но при увеличении количества точек со ста одной до тысячи одной, восстанавливается привычная зависимость, и при дальнейшем росте числа точек норма разности оценённой и заданной функции увеличивается. Видно, что предложенный метод довольно хорошо приближает данное распределение, хотя хуже, чем экспоненциальное распределение, где пять варьируемых параметров достаточно, чтобы получить приближение с ошибкой, меньшей третьего порядка.

Таблица 2 – Ошибка аппроксимации функции плотности гамма распределения

количество точек	dim(**δ**)=1	dim(**δ**)=3	dim(**δ**)=5
11	0.175	0.135	0.006
101	0.16	0.012	0.003
1001	0.253	0.022	0.024

Примечание – функция аппроксимировалась с помощью полиномов Лагерра, приближение проводилось выбором элементов вектора **δ**. Числа в таблице показывают максимум модуля разности между фактическим значением функции плотности распределения и её аппроксимацией среди всех выбранных точек, в которых оценивалась функция плотности распределения. Первый столбец показывает количество точек, в которых оценивалась функция плотности распределения. Остальные столбцы показывают близость приближения в зависимости от количества варьируемых параметров, которые задавались с помощью вектора **δ**.

В таблице 3 представлена норма разности оценённой и заданной функции плотности вероятности распределения хи-квадрат с пятью степенями свободы. Функции рассматривались на отрезке от нуля до десяти. Точки, в которых оценивались функции, получались разбиением данного отрезка на десять, сто и тысячу одинаковых частей, и брались концы полученных отрезков. Аппроксимация хи-квадрат довольно близкая к реальной функции и значительно приближается при увеличении количества варьируемых параметров, точность приближения увеличивается почти на два порядка, при увеличении числа варьируемых параметров с одного до трёх. При увеличении числа варьируемых параметров с трёх до пяти, точность приближения улучшается не на порядки, но всё равно значительно. Ошибка увеличивается с увеличением количества точек, в которых приближается функция, что полностью соответствует ожиданиям. Интересно, что при трёх варьируемых параметрах, максимум разности между оценённой и реальной функциями почти не изменятся при увеличении количества точек, в которых проводится измерение.

Таблица 3 – Ошибка аппроксимации функции плотности хи-квадрат распределения

количество точек	dim(**δ**)=1	dim(**δ**)=3	dim(**δ**)=5
11	0.127	0.006	0
101	0.138	0.007	0.001
1001	0.18	0.007	0.002

Примечание – функция аппроксимировалась с помощью полиномов Лагерра, приближение проводилось выбором элементов вектора **δ**. Числа в таблице показывают максимум модуля разности между фактическим значением функции плотности распределения и её аппроксимацией среди всех выбранных точек, в которых оценивалась функция плотности распределения. Первый столбец показывает количество точек, в которых оценивалась функция плотности распределения. Остальные столбцы показывают близость приближения в зависимости от количества варьируемых параметров, которые задавались с помощью вектора **δ**.

Ошибки аппроксимации бета распределения указаны в таблице 4. Распределение бралось с параметрами два и пять. Функция оценивалась на отрезке от нуля до единицы, точки брались из его равномерного разбиения на количество частей, указанных в первом столбце таблицы. Из таблицы видно, что данное распределение оценивается с большой погрешностью, для точной его оценки требуется значительно больше варьируемых параметров, чем для других рассмотренных распределений. При этом явно отслеживается закономерность, указывающая на то, что при увеличении числа варьируемых параметров, используемых для построения приближения, точность аппроксимации растёт. Также видно, что при увеличении числа точек, в которых рассматривается функция, данная ошибка также

возрастает, что соответствует ожиданием. Другая особенность: видно, что при увеличении числа варьируемых параметров с одного до трёх ошибка может возрастать. Связано это с тем, что для описания данного распределения требуется значительно большее число параметров, и при одном варьируемом параметре сложно говорить о какой-нибудь точности приближения, оценённая функция даже не имеет перегибов, в то время как заданная функция меняет знак монотонности. Только начиная с трёх параметров оценённая функция начинает принимать похожий вид и при дальнейшем увеличении числа варьируемых параметров, точность приближения возрастает.

Таблица 4 – Ошибка аппроксимации функции плотности бета распределения

количество точек	dim(**δ**)=1	dim(**δ**)=3	dim(**δ**)=5
11	1.703	1.539	1.263
101	1.956	2.17	1.849
1001	1.962	2.243	1.921

Примечание – функция аппроксимировалась с помощью полиномов Лагерра, приближение проводилось выбором элементов вектора **δ**. Числа в таблице показывают максимум модуля разности между фактическим значением функции плотности распределения и её аппроксимацией среди всех выбранных точек, в которых оценивалась функция плотности распределения. Первый столбец показывает количество точек, в которых оценивалась функция плотности распределения. Остальные столбцы показывают близость приближения в зависимости от количества варьируемых параметров, которые задавались с помощью вектора **δ**.

Неожиданным результатом явилось то, что простейшее распределение – равномерное на отрезке от нуля до единицы оказалось одним из самых сложных для аппроксимации. В таблице 5 показаны ошибки приближения функции плотности равномерного распределения, которая, в данном случае, является простой константой равной единице. Из неё видно, что точность приближения растёт с увеличением количества варьируемых параметров, но в тоже время рост точности очень медленный, что говорит о том, что для приближения константы необходимо очень много коэффициентов. В тоже время ошибка почти не зависит от количества рассматриваемых точек, что можно объяснить тем, что близость функции к константе не зависит от того, в скольких точках рассматривается функция (стоит напомнить, что близость определяется как супремум модуля разности функций, который, в данном случае, определяется как максимум модуля разности аппроксимированной функции и заранее заданной функции). Чтобы приблизить функции такого рода, лучше всего использовать малое количество точек, в которых рассматривается приближение (чтобы сократить время расчётов), и по возможности

наибольшее количество варьируемых параметров (так как они влияют на точность приближения).

Таблица 5 – Ошибка аппроксимации функции плотности равномерного распределения

количество точек	dim(**δ**)=1	dim(**δ**)=3	dim(**δ**)=5
11	0.638	0.567	0.535
101	0.641	0.57	0.53
1001	0.641	0.57	0.529

Примечание – функция аппроксимировалась с помощью полиномов Лагерра, приближение проводилось выбором элементов вектора **δ**. Числа в таблице показывают максимум модуля разности между фактическим значением функции плотности распределения и её аппроксимацией среди всех выбранных точек, в которых оценивалась функция плотности распределения. Первый столбец показывает количество точек, в которых оценивалась функция плотности распределения. Остальные столбцы показывают близость приближения в зависимости от количества варьируемых параметров, которые задавались с помощью вектора **δ**.

В следующем тесте, результаты которого описаны в таблице 6, рассматривается аппроксимация функции плотности распределения Гумбеля. Данное распределение бралось с параметрами 0, 3. Оно рассматривалось на отрезке от нуля до двадцати. Точки, в которых приближалась данная функция получались с помощью равномерного разбиения данного отрезка на десять, сто и тысячу равных частей (рассматривались концы отрезков, поэтому в первом столбце таблицы присутствуют числа одиннадцать, сто один и тысяча один). Как видно из таблицы, точность приближения значительно улучшается с увеличением числа варьируемых параметров. При увеличении числа варьируемых параметров с одного до трёх, точность растёт от трёх до восьми раз. Также видно, что точность приближения зависит от количества точек, в которых производится измерение. Так, при переходе от одиннадцати точек к сто одной точке, меняется характер поведения точности приближения при увеличении количества варьируемых параметров. При одном варьируемом параметре, при увеличении числа варьируемых параметров с одного до трёх, точность растёт в восемь раз, в то время как при ста одной или тысяча одной точке, при данном переходе точность увеличивается только в три раза.

Таблица 6 – Ошибка аппроксимации функции плотности распределения Гумбеля

количество точек	dim(**δ**)=1	dim(**δ**)=3	dim(**δ**)=5
11	0.148	0.018	0.021
101	0.165	0.049	0.028
1001	0.165	0.048	0.028

Примечание – функция аппроксимировалась с помощью полиномов Лагерра, приближение проводилось выбором элементов вектора **δ**. Числа в таблице показывают максимум модуля разности между фактическим значением функции плотности распределения и её аппроксимацией среди всех выбранных точек, в которых оценивалась функция плотности распределения. Первый столбец показывает количество точек, в которых оценивалась функция плотности распределения. Остальные столбцы показывают близость приближения в зависимости от количества варьируемых параметров, которые задавались с помощью вектора **δ**.

В статье [8] авторы разрабатывают модель равновесия с гетерогенными фирмами, учитывающую идиосинкразические шоки эффективности отдельных фирм. В неё включено понятие неопределённости будущих шоков продуктивности, которые приводят к гистерезису фирм по их экспортному статусу. Компания может войти на внешние рынки, как только она достигнет необходимых для этого размеров и продолжить присутствовать на них, даже если в результате шока продуктивности их деятельность на внешних рынках окажется убыточной. Развитая модель описывает появление фирмы, её рост, изменение её экспортного статуса, а также её закрытия. Модель настраивается с использованием данных по США, таким образом исследуется как изменение размера невозвратных издержек от выхода на экспортные рынки влияет на продуктивность фирм.

В большинстве случаев самые эффективные фирмы участвуют в экспортной деятельности, что может быть связано с тем, что выход на внешние рынки связан с дополнительными издержками. Большинство фирм перестают заниматься экспортной деятельностью в течение первого года после вхождения на иностранные рынки, но те фирмы, что продолжают свою международную деятельность после первого года, процветают. В тоже время, фирмы можно разделить на группы по поведению их продуктивности после выхода на экспортные рынки, многие из них испытывают скачкообразные изменения в размере и эффективности. Но несмотря на гетерогенность фирм, их экспортный статус редко меняется со временем. Присутствие невозвратных издержек оказывает существенное влияние на решение о выходе на экспортные рынки. Во многом, решение принимается, исходя из предыдущего опыта каждой фирмы.

Работа основывается на модели, построенной в [9] с тем отличием, что эффективность предприятий не считается фиксированной и может меняться со временем, согласно стохастическому процессу броуновского движения. Присутствие невозвратных издержек говорит о том, что фирмы несут потери при решении выйти на экспортный рынок. Помимо невозвратных издержек, присутствует неопределённость, касающаяся будущей продуктивности. Предполагается, что при создании фирмы её уровень производительности

определяется случайно, согласно некоторому, одинаковому для всех распределению. Также, создание фирмы связано с невозвратными потерями. После появления фирмы, её уровень производительности плавает со временем согласно броуновскому движению. Фирма может существовать даже при отрицательной прибыли, если её ожидаемая прибыль остаётся положительной. Каждая компания, которая на начале своей деятельности не могла участвовать в экспортной деятельности с положительной прибылью, со временем может стать экспортёром в результате положительного шока производительности.

Если у предприятия с положительным экспортным статусом упала производительность до уровня, при котором фирма не может продолжать получать положительную прибыль от экспортной деятельности, она не останавливает свою экспортную деятельность. Связано это с неопределённостью и невозвратными издержками от выхода на экспорт. Каждый раз при начале экспортной деятельности фирме необходимо понести невозвратные издержки для начала торговли с другими странами, поэтому возможно, что ей выгоднее какое-то время оперировать себе в убыток до тех пор, пока она не испытает очередной шок производительности и её экспортная деятельность вновь начнёт приносить прибыль, и ей не придётся нести повторно невозвратные издержки от выхода на экспорт. Получается, фирмы уходят с экспортной деятельности с некоторой задержкой. Похожая ситуация возникает и перед решением о начале экспортной деятельности: фирмы будут ждать больший промежуток времени для выхода на экспортные рынки, и не будут становиться экспортёрами, даже когда их эффективность достигнет уровня, при котором экспортная деятельность будет приносить положительную прибыль. Из-за неопределённости, связанной с изменением эффективности во времени и присутствием невозвратных издержек, фирме необходимо иметь запас прочности перед началом внешнеэкономической деятельности. Отсюда и возникает гистерезис относительно экспортного статуса предприятий. В пограничных ситуациях компаниям выгоднее не менять свой экспортный статус сразу после шока производительности, поэтому вопрос о том, является предприятие экспортёром или нет, зависит не только от уровня эффективности фирмы, но также и от её экспортного статуса в предыдущие моменты времени.

Рассматривается модель общего равновесия с гетерогенными фирмами, в котором каждая компания принимает решение, выходить на экспорт или нет. Также присутствует неопределённость относительно уровня производительности, так как он меняется со временем согласно броуновскому движению. Из-за данной неопределённости множество фирм с разными экспортными статусами пересекаются на значениях производительности, близких к границе отсечения.

2 Методы оценки структурных эконометрических моделей высоких размерностей

2.1 Общие принципы и подходы

В современной эмпирической экономике всё популярнее становятся данные высокой размерности (данные, где количество доступных переменных велико по сравнению с количеством доступных измерений) [10].

Статья [10] описывает, как можно работать с данными высокой размерности. Подобного рода данные могут возникать из различного рода наблюдений, где каждому объекту присваивают много различных свойств. Например, в американском обследовании доходов и участия в программах собираются сотни характеристик о каждом человеке. Также они могут быть получены автоматически, например, данные о банковских транзакциях, они могут включать в себя тысячи разных переменных, в зависимости от того, какие транзакции хочет выделить исследователь. Вышеперечисленные примеры демонстрируют данные, которые имеют свойство высокой размерности изначально, по своей природе. Но данные высокой размерности могут быть получены и другим путём.

Когда набор данных имеет относительно невысокое число переменных, изначально ничего не известно об их функциональной зависимости между собой. Разного рода функциональные зависимости можно аппроксимировать посредством комбинирования доступных переменных в виде различных функциональных форм (например, их произведения с разными степенями, чтобы получить коэффициенты для ряда Тейлора). Таким образом, даже изначально данные малой размерности могут привести к набору данных с высокой размерностью.

Как показано в книге [11], существует множество статистических методов для построения прогнозов по данным высокой размерности. Но они мало пригодны для задач, где основная цель – оценка параметров модели [12]. Методы интеллектуального анализа данных могут быть изменены и применены к задачам оценки параметров модели по данным высокой размерности. Например, с их помощью можно найти мешающие переменные, присутствующие в данных.

В основном все методы, работающие с данными высокой размерности, сводятся к тому, чтобы каким-то образом понизить их размерность, данный процесс называется регуляризацией. Один из самых простых примеров, показывающий необходимость регуляризации, представляют собой данные, где количество наблюдений совпадает с количеством переменных. В таком случае, метод наименьших квадратов подгонит модель в точности под данные, параметр R^2 будет равен единице. Проблема в том, что помимо реальных данных, она будет учитывать и ненужный шум в данных. Полученная таким образом модель даёт мало информации о реальном процессе. Если в неё подставить другие

данные, то полученные оценки будут далеки от реальных. Чтобы построить по таким данным модель, отражающую реальные процессы, сначала их необходимо регуляризировать.

Среди регрессионных моделей на данных высокой размерности можно выделить их класс, называющийся разреженными моделями на данных высоких размерностей. Данные модели отличаются тем, что среди большого числа потенциально значимых переменных только небольшая их доля оказывает существенное влияние на изучаемую величину. Причём исследователю заранее неизвестно, какие именно переменные дают основной вклад. Одна из возникающих задач - построить прогноз значений исследуемой переменной на других данных. Другая возможная задача заключается в оценке параметров модели на данных высокой размерности. Последняя задача во многом использует похожие методы, как и в задачи построения прогнозов.

Рассмотрим разреженную модель на данных высокой размерности:

$$y_i = g(w_i) + \zeta_i, \tag{15}$$

где y_i – значение исследуемой величины в i-ом наблюдении,

i – индекс наблюдения,

w_i – значение переменных в i-ом наблюдении,

$g()$ – некоторая функция от наблюдаемых величин,

ζ_i – ошибка модели в i-ом наблюдении.

Условное математическое ожидание ошибки ζ в уравнении (15), равно нулю. Также предполагается, что все наблюдения независимы друг от друга. Задача построения модели для прогнозов значений величины y заключается, таким образом, в регуляризации и нахождении функции $g()$.

Не существует какого-то одного единого метода регуляризации данных. Их огромное число. Самым популярным методом регуляризации является отсеивание значимых переменных самим исследователем. Он может основываться на экономической интуиции или на используемой модели. Построенная исследователем модель имеет полезный смысл, только когда в ней находится небольшое число переменных, по которым можно делать дальнейшие выводы. Также предполагается, что все переменные входят в модель в простой функциональной форме, например линейно, что также понижает число возможных

переменных, связанных с взаимодействием различных переменных друг с другом. Но такие методы регуляризации полностью зависят от исследователя, и всегда возникает вопрос, а не ошибся ли он в своих суждениях. Присутствие автоматических методов регуляризации не освобождает исследователя от процесса выбора значимых переменных, но может подсказать на какие переменные стоит обратить внимание.

Как уже говорилось выше, данные высокой размерности могут быть получены даже из небольшого числа переменных, если пытаться построить из них сложную функциональную зависимость. Такого рода проблемы возникают в непараметрических методах (например, [13]). В них предполагается, что искомая функция является достаточно гладкой, чтобы она могла быть аппроксимирована рядом. В них исследователь выбирает значимые для задачи переменные и элементы аппроксимационного ряда. Соответственно, в них главная часть регуляризации, а именно, отбор небольшого числа необходимых переменных и отбор конкретных членов ряда, также лежит на исследователе.

Рассмотрим регуляризацию задачи, где функция $g(w_i)$ может быть представленна в виде:

$$g(w_i) = \sum_{j=1}^{p} \beta_j \, x_{i,j} + r_{p,i}, \tag{16}$$

где $g()$ – функция, описывающая связь между наблюдаемыми переменными и исследуемой величиной,

i – индекс наблюдения,

w_i – значение переменных в i-ом наблюдении,

p – количество переменных в данных,

j – индекс переменной,

β_j – значение j-го коэффициента в линейной регрессии,

$x_{i,j}$ – значение j-ой переменной в i-ом наблюдении,

$r_{p,i}$ – ошибка аппроксимации в i-ом наблюдении.

Функция $g(w_i)$ в выражении (16) не обязательно является линейной. Вектор $x = (x_{i,1}, \dots, x_{i,p})'$ помимо неависимых компонет, может содержать их всевозможные преобразования, чтобы представить функцию $g(w_i)$ в виде некоторого разложения. Общее количество переменных p может быть больше, чем размер выборки n. Ошибка

аппроксимации $r_{p,i}$ считается незначительной, по сравнению с ошибкой в данных ζ_i. Решать напрямую данную модель не получится, так как количество переменных p больше, чем число наблюдений n. Для получения каких-либо выводов из построенной модели её необходимо сначала регуляризовать.

Для регуляризации задачи предполагается, что данная модель является разреженной моделью на данных высокой размерности. Что означает, что только небольшое число s переменных вносят существенный вклад в значение изучаемой величины, или, что то же самое, большая часть коэффициентов β_j равны или очень близки к нулю. В такого рода моделях, значительная часть коэффициентов может быть отброшена, притом, что ошибка аппроксимации $r_{p,i}$ заметно не вырастет. Задача состоит из двух частей: найти коэффициенты, вносящие существенный вклад в исследуемую величину, а потом оценить отобранные коэффициенты.

Разряженные модели на данных высокой размерности включают в себя как обычные параметрические, так и не параметрические модели. Она позволяет объединить оба этих подхода и позволить исследователю рассматривать большое количество переменных, а также большое количество преобразований от них. Она автоматически отбирает значимые коэффициенты, позволяя отобрать небольшое количество переменных, оказывающих существенное влияние на исследуемую величину. Данный метод может быть применён к задачам, где заранее неизвестно, какие переменные являются важными, а какие нет.

Метод для работы с разреженными данными, называемый LASSO, был представлен в работах [14]. [15]. Идея данного метода заключается в том, что коэффициенты выбираются таким образом, чтобы минимизировать серднеквадратичную ошибку плюс штраф за количество используемых переменных. Данный метод представляет собой модифицированный метод наименьших квадратов. Рассмотрим вариант метода LASSO, предложенный в работе [16], где коэффициенты $\hat{\beta}$ представленны в виде:

$$\hat{\beta} = \underset{b}{\operatorname{argmin}} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{i,j}\, b_j \right)^2 + \lambda \sum_{j=1}^{p} |\, b_j | \gamma_j, \tag{17}$$

где $\hat{\beta}$ – оценки коэффициентов модели,

b – вектор коэффициентов в изначальной модели,

i – индекс наблюдения,

y_i – значение исследуемой величины в i-ом наблюдении,

p – количество переменных в данных,

j – индекс переменной,

$x_{i,j}$ – значение j-ой переменной в i-ом наблюдении,

b_j – j-ая компонента вектора b,

λ – коэффициент штрафа за дополнительную переменную,

γ_j – поправочный коэффициент на размерность величины в данных.

В формуле (17) первая часть полностью повторяет метод наименьших квадратов, основное отличие заключается во втором слагаемом. В нём, коэффициент $\lambda > 0$ отвечает за штраф, накладываемый на количество переменных, с помощью него можно управлять силой ограничения на число отобранных переменных, чем он больше, тем меньше количество отбираемых переменных (точнее тем ниже значении коэффициентов b_j у всех переменных, но все значения b_j близкие к нулю отсеиваются). Коэффициент γ необходим для того, чтобы учесть разнородность данных. Одни величины могут выражаться миллионами, в то время как другие единицами. Соответственно некоторые переменные будут иметь большие коэффициенты при них в силу своей природы. Но необходимо учитывать именно вес каждой переменной в силе влияния на исследуемую величину, независимо от её размерности. Данный коэффициент решает эту проблему.

Проблема, решаемая методом LASSO, является выпуклой. Поэтому для неё можно использовать методы выпуклой оптимизации, для которых написаны быстрые алгоритмы решения таких проблем. Так как штрафная функция (17) имеет излом в нуле, то в процессе оптимизации многие коэффициенты станут в точности равны нулю. В некоторых случаях процесс отбора значимых коэффициентов может свестись к отбору ненулевых элементов. Данный метод может применяться в задачах, учитывающих гетероскедастичность данных и ненормальное распределение ошибок [17].

Коэффициенты, полученные методом LASSO, как правило, являются заниженными. Чтобы это исправить, применяется процедура post-LASSO. Заключающаяся в обычной регрессии, которая использует только те переменные, которые были отобраны на первом шаге процедурой LASSO.

Простое использование разреженной модели позволяет решать задачи прогнозирования и отбора переменных, на которые стоит обращать внимание. Но использование полученных коэффициентов как оценку параметров модели показывает не лучший результат. Данная

процедура была разработана для задач построения прогнозов. Чтобы получить выводы о причинно-следственных связях в виде оценки параметров модели, лучше использовать другие методы, которые жертвуют точностью прогнозов в счёт более верных оценок параметров.

В реальности проблема заключается в том, что возможны ошибки отбора нужных переменных. Если предположить, что у нас есть идеальная процедура, которая всегда отбирает необходимые переменные, то с её помощью всегда можно было бы отобрать переменные, и потом на них применять любой удобный метод, который позволит построить нужную модель. Но данный подход не может иметь место в реальности, потому что действительно могут возникать ошибки при выборе переменных, и необходимые переменные не попадут во множество, по которому будет строиться выбранная модель. Данная проблема может возникать при любом методе регуляризации, когда исследователь сам выбирает весомые переменные, он может ошибиться и опустить ряд необходимых переменных. Таким образом, это не только проблема задач на данных с высокой размерностью. Соответственно, желательно создать процедуру, которая по возможности более точно выбирала бы необходимые переменные. В реальных условиях в любом случае используется отбор переменных, процедура, которая позволит отбирать их более точно, помогает избежать ошибок отбора и увеличить точность построенных моделей.

Для получения более робастных процедур используют методы, где исследователь выбирает небольшое количество основных переменных модели, которые входят в модель и не стоит вопроса об их исключении через процедуру отбора. Но отбор среди оставшихся переменных, которые могут влиять на исследуемую величину, проводится с помощью автоматической процедуры. Оценка отобранных основных параметров, проводится с помощью методов, оценки которых не изменяются при добавлении остальных косвенных переменных [18].

Например, в модели, где присутствует большое количество инструментальных переменных:

$$\begin{aligned} y_i &= \alpha d_i + \varepsilon_i \\ d_i &= z_i' \Pi + r_i + v_i, \end{aligned} \tag{18}$$

где i – индекс наблюдения,

y_i – значение исследуемой величины в i-ом наблюдении,

α – коэффициент перед переменной, представляющей интерес,

ε_i – ошибка измерений в i-ом наблюдении,

d_i – значение переменной, представляющей интерес, в i-ом наблюдении,

z_i' – вектор строка значений инструментальных переменных в i-ом наблюдении, входящих в переменную, представляющую интерес,

Π – вектор коэффициентов перед инструментальными переменными в выражении для переменной, представляющей интерес,

r_i – значение ошибки аппроксимации в i-ом наблюдении,

v_i – ошибка измерений в i-ом наблюдении в выражении для переменной, представляющей интерес.

В формуле (18) ошибка ε_i показывает, насколько в модели отклоняется предсказанное значение от реального, используя переменную d как основной показатель, от которого зависит y. Предполагается, что $E[\varepsilon_i|z_i] = 0$. Но переменная d, в свою очередь, может зависеть от переменных z. В линейной функциональной форме, которая аппроксимирует её с ошибкой r. Ошибка v показывает, на сколько предсказанное значение d отличается от действительного. Предполагается, что $E[v_i|z_i, r_i] = 0$. Таким образом, d – скалярная эндогенная переменная, размерность вектора инструментальных переменных p может превышать число доступных наблюдений n. Данную формулу можно обобщить и на большее число эндогенных переменных, если предполагать, что они находятся в ошибке ε, тогда последовательно можно исключать каждую переменную с помощью $y - \alpha d_i = \varepsilon_i = \alpha_1 d_{1,i} + \varepsilon_{1,i}$.

Так как количество возможных инструментальных переменных может быть больше, чем число доступных наблюдений, для построения оценки необходимо провести регуляризацию задачи. Данную задачу можно решить используются разреженные модели. После процедуры отбора инструментальных переменных их останется небольшое число, и по ним можно строить оценки для параметра α с помощью двухэтапного метода наименьших квадратов. Данная оценка использует алгоритм, описанный выше. Сначала выбираются основные переменные, которые обязательно входят в модель, в данном случае – это переменная d. На следующем шаге отбираются косвенные переменные, которые тоже могут влиять на результирующую модель, но в значительно меньшей степени. Они отбираются с помощью процедуры LASSO. На втором шаге двухэтапного метода наименьших квадратов возможные ошибки в процедуре выбора инструментальных переменных не вызывают сильных отклонений в параметрах модели. Соответственно, в

задачах такого типа можно использовать разряженные модели на данных высокой размерности.

Другой тип задач, где может быть применена процедура отбора, представляет собой обычную задачу регрессии с широким набором контрольных переменных. Её можно записать в виде:

$$y_i = \alpha d_i + x_i'\theta_y + r_{yi} + \zeta_i, \tag{19}$$

где i – индекс наблюдения,

y_i – значение исследуемой величины в i-ом наблюдении,

α – коэффициент перед переменной, представляющей интерес,

d_i – значение переменной, представляющей интерес, в i-ом наблюдении,

x_i' – вектор-строка значений контрольных переменных в в i-ом наблюдении,

θ_y – вектор коэффициентов в линейной регрессии,

r_{yi} – значение ошибки аппроксимации в i-ом наблюдении,

ζ_i – ошибка модели в i-ом наблюдении.

Уравнение (19) представляет собой обычную линейную регрессию. Только количество контрольных переменных n может быть значительно больше, чем число наблюдений n. Предполагается, что $E[\zeta_i|d_i, x_i, r_{yi}] = 0$. Переменная, эффект от которой представляет интерес, обозначена как d, остальные регрессоры должны выбираться из различных соображений. Так как их количество велико, то для построения оценок необходимо провести регуляризацию задачи.

Чтобы провести регуляризацию задачи, представленной в уравнении (19), можно предположить, что достаточно применить процедуру LASSO, исключив из неё переменную d, и по отобранным переменным найти коэффициенты модели, с помощью обычного метода наименьших квадратов. Но такой алгоритм приведёт к ошибкам, что не совсем на первый взгляд очевидно. Дело в том, что изначально, метод LASSO разрабатывался для задач построения прогнозов, и, соответственно, если какие-то из двух переменных оказываются сильно коррелированы между собой, одна из них автоматически отбрасывается, так как вторая не несёт много информации для построения прогноза. Но в случае, если стоит задача оценки коэффициента эффекта воздействия α, данный подход

приведёт к серьёзным ошибкам из-за пропущенной переменной. Возникновение пропущенной переменной в такой процедуре неизбежно, если в векторе x присутствуют контрольные переменные, влияющие на переменную d, эффект от которой стоит задача измерить.

Поэтому прямое применение процедуры LASSO к данной задаче неуместно. Из вышесказанного можно понять, что данная процедура не учитывает возможные зависимости между переменной, эффект от который представляет интерес и остальными регрессорами. Данную проблему можно разрешить, если рассмотреть дополнительную задачу, в которой явно ищется зависимость между переменной, эффект от который представляет интерес d и остальными регрессорами x:

$$d_i = x_i'\theta_d + r_{di} + v_i, \tag{20}$$

где i – индекс наблюдения,

d_i – значение переменной, представляющей интерес, в i-ом наблюдении,

x_i' – вектор-строка значений контрольных переменных в в i-ом наблюдении,

θ_d – вектор коэффициентов в сокращённой регрессии,

r_{di} – значение ошибки аппроксимации в i-ом наблюдении, в сокращённой регрессии,

v_i – ошибка сокращённой регрессии в i-ом наблюдении.

Уравнение (20) показывает связь между переменной, эффект которой представляет интерес, и остальными регрессорами. Предполагается, что $E[v_i|x_i, r_{di}]$. С помощью данной формулы можно учесть возможные пропущенные переменные, если бы метод LASSO применялся просто напрямую к изначальному уравнению.

Для того, чтобы получить модель, близко отражающую функциональную зависимость между исследуемой переменной и переменной, эффект которой представляет интерес, изначальное уравнение полезно переписать в редуцированную форму, подставив уравнение (20) в (19). Тогда получится система уравнений:

$$\begin{aligned} y_i &= x_i'(\alpha\theta_d + \theta_y) + (\alpha r_{di} + r_{yi}) + (\alpha v_i + \zeta_i) = x_i'\pi + r_{ci} + \varepsilon_i \\ d_i &= x_i'\theta_d + r_{di} + v_i, \end{aligned} \tag{21}$$

где i – индекс наблюдения,

y_i – значение исследуемой величины в i-ом наблюдении,

$x_i{}'$ – вектор-строка значений контрольных переменных в в i-ом наблюдении,

α – коэффициент перед переменной, представляющей интерес,

θ_d – вектор коэффициентов в сокращённой регрессии,

θ_y – вектор коэффициентов в линейной регрессии,

r_{di} – значение ошибки аппроксимации в i-ом наблюдении, в сокращённой регрессии,

r_{yi} – значение ошибки аппроксимации в i-ом наблюдении,

v_i – ошибка сокращённой регрессии в i-ом наблюдении

ζ_i – ошибка модели в i-ом наблюдении,

π – вектор коэффициентов регрессии в редуцированной форме,

r_{ci} – значение суммарной ошибки аппроксимации в i-ом наблюдении,

ε_i – значение суммарной ошибки моделей в i-ом наблюдении,

d_i – значение переменной, представляющей интерес, в i-ом наблюдении.

Уравнение (21) представляет собой редуцированную форму изначальной задачи, в нём $E[v_i|x_i, r_{di}] = 0$. Оба данных выражения представляют собой некоторую модель, которая может быть оценена с использованием разреженных моделей на данных высокой размерности.

К любому из данных уравнений может быть применена процедура отбора. На самом деле это является необходимым условием, так как в них количество переменных может превышать количество наблюдений. Также, если отсеивать переменные только по одному из представленных выражений, будет получена ошибка, связанная с возможными пропущенными переменными. В предположении, что процедура отбора безошибочна и всегда выдаёт все нужные переменные, было бы достаточно отобрать все переменные из первого уравнения. Но в реальности, как уже говорилось, исследователи могут ошибаться и отбирать не все нужные переменные. Если взять первое уравнение, где описывается зависимость y от x, процедура отбора отсеет все переменные из x, которые оказывают малое влияние на y. Но те же самые переменные могут оказывать сильное воздействие на d, тогда не включение данных характеристик наблюдений приведёт к построению модели, далёкой от реальности. В ней явно будут присутствовать пропущенные переменные. Похожая ситуация произойдёт, если отсеивать только по второму уравнению, описывающему зависимость d от x. Тогда могут быть выброшены переменные, слабо

влияющие на d, но имеющие сильное воздействие на y, что также приведёт к пропущенным переменным.

Чтобы учесть возможные ошибки такого рода, необходимо использовать оба представленных уравнения. Тогда процедура отбора переменных будет проходить в два этапа: сначала отбираются переменные из первого уравнения, потом независимо от первого шага отбираются переменные из второго уравнения, и в итоге берётся объединение отобранных на двух шагах переменных. После на отобранных переменных проводится процедура оценивания коэффициента α. Используя два шага, отбирают переменные, которые оказывают сильное влияние как на y, так и на d. Когда берётся объединение отобранных переменных из двух уравнений, вероятность получения ошибки, вызванной пропущенной переменной, существенно снижается, потому что используются как переменные, связанные с y, так и переменные, связанные с d.

Описанная двушаговая процедура отбора исключает ошибки, возникающие при использовании процедуры отбора только на одном из уравнений. По крайней мере используются переменные, которые имеют сильное воздействие как на y, так и на d. Соответственно, отброшенные переменные имеют незначительное влияние как на y, так и на d, что ограничивает влияние возможных пропущенных переменных. Процедура двойного отбора неявно проводит регрессию ошибки ε_i на v_i, для построения оценки α, что похоже на метод предложенный в [19]. Данная процедура успешнее избавляет от пропущенных переменных, чем прямой метод отбора на симулированных данных.

В статье [10] был рассмотрен пример применения процедуры двойного выбора к задаче оценки эффекта воздействия закона о легализации абортов на уровень преступности. Они основывались на работе по данной теме [20], согласно которой, по данным на уровне страны увеличенное количество абортов в 70-х годах привело к понижению уровня преступности двумя декадами позже. Но в подобного рода данных присутствует эндогенность, а законы, принимаемые каждым отдельным штатом, не были выбраны случайным образом. Это осложняет процесс выявления причинно-следственных связей, так как возможны присутствия переменных, которые влияют одновременно на оба изучаемых параметра. Если не включить их в список контрольных переменных, то полученная оценка может быть смещённой из-за присутствия пропущенной переменной. Чтобы учесть данную проблему, авторы [20] использовали метод разницы в различиях, их модель выглядела следующим образом:

$$y_{cit} = \alpha_c a_{cit} + w_{it}{}'\beta_c + \delta_{ci} + \gamma_{ct} + \varepsilon_{cit}, \qquad (22)$$

где c – индекс, указывающий на тип преступления,

i – индекс, показывающий штат, в котором бралось наблюдение,

t – год измерения,

y_{cit} – количество преступления типа c, в штате i, в год t,

α_c – коэффициент влияния количества абортов, на количество преступлений типа c,

a_{cit} – количество абортов в штате i, в год t, связанных с типом преступлений c (определяется по возрасту),

$w_{it}{}'$ – вектор-строка доступных контрольных переменных для штата i, в год t,

β_c – регрессионные коэффициенты влияния контрольных переменных на уровень преступности относительно преступлений типа c,

δ_{ci} – эффект, влияющий на уровень преступлений типа c, в штате i, специфичный для данного штата, не меняющийся со временем,

γ_{ct} – тренд национального уровня, меняющийся во времени, относительно количества преступлений типа c, в момент времени t,

ε_{cit} – ошибка модели при оценке уровня преступности типа c, в штате i, в год t.

Уравнение (22) представляет собой линейную регрессионную модель, выделяющую в отдельную переменную количество абортов и рассматривающая остальные доступные переменные как контрольные. Среди контрольных переменных присутствовали такие переменные, как уровень безработицы, уровень бедности, уровень потребления алкоголя и другие. Оценка данной модели на изучаемый эффект говорит о существенном влиянии легализации абортов на уменьшение уровня преступности. Например, увеличение числа абортов до 100 на 1000 рождённых детей уменьшает уровень преступности на 15%. Данный вывод можно считать верным, если предположить, что все потенциальные факторы, не включённые в вектор w_{it} заключены в национальные тренды или не меняются со временем.

Добавление в модель переменных национального тренда и факторов, не меняющихся во времени, позволяет говорить о корректных выводах во многих различных ситуациях. Они добавляют модели гибкости. Но всё же, можно говорить о корректности полученных выводов только в том случае, если все изменяющиеся во времени переменные, особенные для каждого штата, коррелированные как с уровнем абортов, так и уровнем преступности, присутствуют в векторе w_{it}. Один из способов расширить поставленную модель – добавить в неё набор линейных трендов, специфичных для каждого штата в отдельности. Только

данный подход добавляет большое количество переменных в модель. Также, линейная аппроксимация приближает функцию только на небольших интервалах её изменений, что не позволяет строить выводы о долгосрочных данных.

Методы, использующие данные высокой размерности, позволяют вводить дополнительные переменные в большом количестве, независимо от числа доступных наблюдений. Чтобы учесть пропущенные переменные, можно добавить в модель нелинейные тренды, влияющие на отобранные контрольные переменные, на уровень преступности и на уровень количества абортов. С одной стороны, данный подход позволяет учесть многие не включенные эффекты в модель, с другой стороны, он добавляет в модель огромное число дополнительных переменных, что не позволяет его использовать для оценки обычными методами. Но данную задачу можно регуляризировать с помощью отбора переменных и, таким образом, оставлять только те добавленные факторы, которые сильно влияют на изучаемое явление. В данном случае идея заключалась в том, чтобы для каждого штата построить нелинейный тренд, влияющий как на число абортов, так и на уровень преступности, чтобы с помощью него учесть эффекты от не наблюдаемых переменных. Для этого рассматривалась модель в виде:

$$\begin{aligned} \Delta y_{cit} &= \alpha_c \Delta a_{cit} + z_{cit}{}' + \overline{\gamma}_{ct} + \Delta \varepsilon_{cit} \\ \Delta a_{cit} &= z_{cit}{}' \Pi_c + \overline{\kappa}_{ct} + \Delta v_{cit}, n \end{aligned} \tag{23}$$

i – индекс, показывающий штат, в котором бралось наблюдение,

t – год измерения,

Δy_{cit} – изменение количества преступлений типа c, в штате i, от года $t-1$ к году t,

α_c – коэффициент влияния количества абортов, на количество преступлений типа c,

Δa_{cit} – изменение количества абортов в штате i, связанных с типом преступлений c (определяется по возрасту), от года $t-1$ к году t,

$z_{cit}{}'$ – вектор-строка широкого набора контрольных переменных для штата i, в год t, влияющих на уровень преступления типа c,

β_c – регрессионные коэффициенты влияния контрольных переменных на уровень преступности относительно преступлений типа c,

$\overline{\gamma}_{ct}$ – эффект, меняющийся во времени, влияющий на изменение количества преступлений типа c, от года $t-1$ к году t,

$\Delta\varepsilon_{cit}$ – изменение ошибки модели при оценке уровня преступности типа c, в штате i, от года $t-1$ к году t,

Π_c – вектор коэффициентов, влияющих на изменение количества абортов,

$\overline{\kappa}_{ct}$ –эффект, меняющийся во времени, влияющий на изменение количества абортов, от года $t-1$ к году t,

Δv_{cit} – зменение ошибки модели при оценке количества абортов, в штате i, от года $t-1$ к году t,

В формуле (23), используется изменение уровня преступности $\Delta y_{cit} = y_{cit} - y_{cit-1}$. Данное уравнение использует набор инструментальных переменных z_{cit}. Их количество может превышать число наблюдений. Она состоит из переменных, доступных в данных, а также из их различных взаимодействий дру с другом до третьего порядка включительно (произведения переменных друг на друга и самих на себя, где суммарная степень меньше либо равна трём). Но в данном конкретном случае количество переменных было равно 284, в то время как число доступных наблюдений равнялось 600. Таким образом, в данной задаче было возможно найти коэффициенты обычной, прямой регрессии. Но такая задача не будет давать надёжные результаты. В данном случае, оценённый коэффициент α является положительным равным 0.07, что соответсвует увеличению количества преступлений. Доверительный интервал данной оценки в 10 раз шире, полученного ранее значения. Такой результат сильно отличается от оценок, полученных другими методами.

Процедура двойного выбора отбирает значительные коэффициенты по отдельности из каждого уравнения. Если к данной модели применить процедуру двойного выбора, то полученные оценки коэффициентов получаются близки к оценкам, полученными изначально. Только значительно увеличивается 95% доверительный интервал. В полученном доверительном интервале будут присутствовать как очень большие отрицательные значения коэффициента α, так и его положительные значения. Но данный результат значительно точнее, чем оценки, полученные моделью, в которую включены все переменные – по сравнению с ней доверительный интервал в три раза меньше.

Получили, что оценки с помощью метода двойного отбора могут качественно отличаться от результатов, полученных с помощью регрессии, где было выбрано небольшое число значимых переменных. Что говорит о том, что выводы, сделанные на основе линейной регрессии недостаточно робастные, чтобы однозначно говорить о виде причинно-следственной связи между легализацией абортов и уровнем преступности.

Другой интересный пример, рассмотренный в [16], основан на работе [21] где оценивается влияние институциональной среды на общий выпуск в экономике. Данная проблема усложняется тем, что институциональная среда может влиять не только на выпуск, но и, например, на доход населения, а повышенный доход, в свою очередь, приводит к улучшению институтов. Чтобы обойти данную проблему использовался показатель смертности среди первых колонистов как инструментальная переменная для показателя развитости институтов. Данное решение основано на предположении, что долгосрочные поселения с большой вероятностью появятся в тех местах, где была более развитая институциональная среда. Таким образом, уровень смертности на начальном этапе колонизации коррелирован с развитием институтов в данных поселениях. В дальнейшем предполагалось, что уровень экономического выпуска сотни лет спустя мало зависит от уровня смертности в прошлых веках, за исключением тех случаев, кода имелись заложенные институты на момент основания. С помощью такого подхода оценивались влияния развития институтов на уровень выпуска в экономике.

Чтобы модель, построенная на основе инструментальных переменных была верна, необходимо, чтобы другие рассмотренные факторы, влияющие на валовый внутренний продукт, были инварианты во времени. Таким фактором может выступить географическое положение поселения. Географическими данными может быть удалённость от экватора, а также принадлежность к какому-либо континенту. Исследуемую величину можно контролировать, используя географические данные. Используя доступные данные, можно построить модель:

$$\log(GDPpp_i) = \alpha \cdot ExpropProt_i + x_i'\beta + \varepsilon_i, \tag{24}$$

где i – индекс, указывающий наблюдаемую страну,

$GDPpp_i$ – ВВП на душу населения страны i,

α – коэффициент воздействия защиты прав собственности на уровень выпуска,

$ExpropProt_i$ – уровень защиты прав собстенности в стране i,

x_i' – вектор-строка доступных контрольных переменных,

β – вектор коэффициентов регрессии для контрольных переменных,

ε_i – ошибка модели в i-ом измерении.

В формуле (24) представлена модель, по которой оценивалось влияние институциональной среды на выпуск экономики. Вектор контрольных переменных x_i представляет собой набор

географических показателей стран. $ExpropProt_i$ показывает меру того, насколько сильно в стране i защищены права собственности. Она используется как прокси переменная для показателя развитости институциональной среды в стране i. Предполагается, что уровень смертности в начальном поселении является корректной инструментальной переменной при контроле географических данных. Можно рассматривать контрольные переменные в виде набора дамми переменных и линейной зависимости от дальности от экватора. Но использование методов для работы с данными большой размерности позволяет расширить рассмотрение данной зависимости. С помощью отбора переменных, связанных с изучаемым явлением не обязательно в линейном виде.

Но чтобы использовать методы высокой размерности, необходимо учесть, что при явном их использовании могут возникать ошибки, связанные с пропущенной переменной. Чтобы их устранить, необходимо явно выписать переменные, которые могут быть потенциально эндогенными и отбирать переменные в каждом из полученных уравнений. Тогда получим систему:

$$\begin{aligned} \log(GDPpp_i) &= \alpha \cdot ExpropProt_i + x_i'\beta + \varepsilon_i \\ ExpropProt_i &= \pi_1 \cdot SlrMort_i + x_i'\Pi_2 + v_i \\ SlrMort_i &= x_i'\gamma + u_i, \end{aligned} \tag{25}$$

где i – индекс, указывающий наблюдаемую страну,

$GDPpp_i$ – ВВП на душу населения страны i,

α – коэффициент воздействия защиты прав собственности на уровень выпуска,

$ExpropProt_i$ – уровень защиты прав собстенности в стране i,

x_i' – вектор-строка доступных контрольных переменных,

β – вектор коэффициентов регрессии для контрольных переменных,

ε_i – ошибка модели в i-ом измерении,

π_1 – коэффициент зависимости между уровнем защиты собственности и уровнем смертности среди первых поселенцев,

$SlrMort_i$ – уровень смертности первых поселенцев в стране i,

Π_2 – вектор коэффициентов регрессии, для контрольных переменных в модели для уровня защиты собственности,

v_i – ошибка модели уровня защиты собственности в i-ом измерении,

γ – вектор коэффициентов регрессии, для контрольных переменных в модели для уровня смертности среди первых поселенцев,

u_i – ошибка модели для уровня смертности среди первых поселенцев в i-ом измерении.

Уравнение (25) можно привести в редуцированную форму в виде:

$$\begin{aligned} \log(GDPpp_i) &= x_i'\tilde{\beta} + \tilde{\varepsilon}_i \\ ExpropProt_i &= x_i'\tilde{\Pi}_2 + \tilde{v}_i \\ SlrMort_i &= x_i'\gamma + u_i, \end{aligned} \tag{26}$$

где i – индекс, указывающий наблюдаемую страну,

$GDPpp_i$ – ВВП на душу населения страны i,

x_i' – вектор-строка доступных контрольных переменных,

$\tilde{\beta}$ – вектор коэффициентов редуцированной регрессии для контрольных переменных,

$ExpropProt_i$ – уровень защиты прав собстенности в стране i,

$\tilde{\Pi}_2$ – вектор коэффициентов редуцированной регрессии, для контрольных переменных в модели для уровня защиты собственности,

$\tilde{v}_i$ – ошибка редуцированной модели уровня защиты собственности в i-ом измерении,

$SlrMort_i$ – уровень смертности первых поселенцев в стране i, γ – вектор коэффициентов регрессии, для контрольных переменных в модели для уровня смертности среди первых поселенцев,

u_i – ошибка модели для уровня смертности среди первых поселенцев в i-ом измерении.

Уравнение (26) представляет собой редуцированную форму регрессий. В данной системе присутствуют только географические переменные. К ней можно применить процедуру отбора переменных, оказывающую существенное влияние на исследуемые величины. Необходимо применить процедуру LASSO к каждому из трёх уравнений. Потому нужно взять объединение всех переменных, отобранных из каждого уравнения. Полученное множество факторов, можно использовать для оценивания параметров обычными методами, применяемыми для уравнений с инструментальными переменными. В данном случае переменная, показывающая уровень смертности среди первых поселенцев, является инструментальной переменной для уровня защиты прав собственности, которая, в свою очередь, является прокси переменной для уровня развития институциональной среды.

Важно отметить, что в данном случае основные переменные выбирались исследователем. Метод двойного выбора использовался только для того, чтобы отобрать дополнительные переменные, которые могут влиять на исследуемые характеристики. Полученные результаты во многом совпадают с теми, что были извлечены из модели, учитывающей только линейное расстояние от экватора и дамми переменные на континент. Согласно оценке, полученной методом двойного отбора, сильное положительное влияние институциональной среды на выпуск в экономике действительно присутствует, хоть оценка и получилось ниже, чем при простой линейной регрессии. В то же время, доверительный интервал полученной оценки немного понизился.

Таким образом, метод двойного отбора позволяет использовать более гибкие функции и различные переменные, отбирая из них небольшое число параметров, которые влияют на переменные, представляющие интерес. Таким образом методы высоких размерностей могут дополнить набор инструментов, доступный исследователям. Разреженные модели на данных высокой размерности позволяют использовать метод автоматической регуляризации данных, для того чтобы к ним можно было применять другие методы исследования.

2.2 Сравнительный анализ методов

В эконометрической литературе существует большое количество различных методов, пригодных для оценки эффектов воздействия. К сожалению, большинство из них не вполне применимо в ситуациях, которые характеризуются широким набором контрольных переменных, в то время как именно эта особенность явно рассматривается в нашей работе. Соответственно, мы останавливаемся на рассмотрении только тех методов, которые явным образом пригодны для подобных случаев, что означает, что основной выбор происходит между двумя основными пластами литературы, существенно связанными между собой.

Первое направление в литературе восходит к исходной идее двойного выбора, предложенного в работе [22], которая, впоследствии была существенно обобщена и расширена. Ключевая идея данного подхода заключается в построении таких эконометрических моделей и процедур, чтобы ошибки в оценке бесконечномерного вторичного параметра не слишком сильно влияли на результаты оценивания первичного параметра. Формальное изложение этого общего подхода изложено, например, в статье [23]. Технически для успешной применимости метода требуется, чтобы целевая функция рассматриваемой оценки имела производную Гато по бесконечномерному мешающему параметру равную нулю. На практике это означает что, с одной стороны, данный подход может быть успешно применен к широкому набору методов оценивания и моделей, однако с другой стороны, он требует достаточно высокой степени технической и математической

квалификации для его применения в нестандартных ситуациях. Соответственно, на практике чаще всего применяются известные частные случаи, в первую очередь двойной выбор, хотя более сложные методы и продолжают активно набирать популярность.

К ключевым преимуществам данного подхода можно отнести, в первую очередь, его близость к классическим эконометрическим методам и подходам, что позволяет использовать его в широком наборе привычных контекстов. Кроме того, это существенно облегчает интерпретацию полученных результатов. С другой стороны, это может служить и существенным недостатком, поскольку может приводить к тому, что методы часто применяются в тех ситуациях, когда требуемые предположения не выполнены, например когда не реалистично предположение о разреженности или необходимы предположения о структуре причинно-следственных связей между рассматриваемыми переменными. Однако на практике такие же проблемы существуют у большинства эконометрических методов.

Другая группа подходов основана на применении ансамблей деревьев, например, случайных лесов, и идеи честного оценивания. Исходно идея была предложена в работе [24] и доработана в статье [25].

При использовании деревьев для оценки неоднородного эффекта воздействия все наблюдения разбиваются в неоднородные группы в зависимости от значений контрольных переменных. В рамках каждой группы эффект воздействия оценивается одинаково, посредством сравнения исходных внутри одной группы. Данный подход, в целом, близок к традиционному сопоставлению по мере склонности, однако за счет свойств ансамблевых методов данный подход остается валидным в условиях существования большого количества контрольных переменных. При этом не всякий подход дает возможность получить валидные доверительные интервалы и тесты на значимость эффекта воздействия. Для того, чтобы гарантировать их валидность, необходимо, чтобы рассматриваемая процедура являлась “честной”, в том смысле, что каждое наблюдение может использоваться в дереве либо для выбора того, к какому из листьев следует отнести заданное наблюдение, либо для определения требуемого эффекта внутри листа.

3 Оценка роли эффекта отбора в различиях между экспортирующими и неэкспортирующими предприятиями в России

3.3 Основные результаты

Основываясь на полученных оценках меры эффективности фирм, изучаются влияние импорта и экспорта компаний на их производительность.

В таблице 7 в строках используются следующие обозначения: «Константа» показывает значение константы в регрессии, «Импорт» – коэффициент, перед дамми переменной, являющейся индикатором импортной активности компании, он отражает уровень влияния импортной деятельности на производительность. «Экспорт», как и «Импорт», показывает значение коэффициента перед дамми переменной, являющейся индикатором экспортной деятельности, данный коэффициент характеризует меру влияния участия в экспортной деятельности на производительность; переменная «Импорт*Экспорт» показывает значение коэффициента перед произведением индикаторов участия в экспортной и импортной деятельностей, она отражает силу нелинейного совместного взаимодействия импортной и экспортной деятельности фирмы на её уровень эффективности.

Из таблицы 7 видно, что все оценённые коэффициенты являются статистически значимыми. При изменении модели, результаты качественно не изменяются. Так же можно заметить, что при добавлении нелинейных слагаемых в регрессию, коэффициенты, описывающие влияние экспортного статуса на эффективность фирм, уменьшаются. Что даёт основание предполагать существующую нелинейную зависимость между экспортным статусом фирмы и изменением её уровня производительности.

Таблица 7 – Значение оценки коэффициентов в регрессии эффективности фирмы на индикатор внешнеэкономической деятельности

	Модель1 (экспорт)	Модель2 (импорт)	Модель3 (экспорт* импорт)	Модель4 (экспорт + импорт)	Модель5 (экспорт + импорт + экспорт*импорт)
Константа	38.218*** (0.013)	41.222*** (0.035)	37.321*** (0.115)	36.572*** (0.121)	39.001*** (0.113)
Экспорт	6.462*** (0.115)	-	-	2.15*** (0.127)	1.23*** (0.21)
Импорт	-	7.028*** (0.092)	-	4.358*** (0.095)	3.318*** (0.142)
Импорт*Экспорт	-	-	7.231*** (0.132)	-	5.126*** (0.193)

Примечание: результаты получены с помощью метода наименьших квадратов с использованием предложенной в работе меры эффективности в качестве зависимой переменной. В скобках указаны стандартные ошибки оценок, устойчивые к гетероскедастичности и кластеризации ошибок. Звездочками отмечены коэффициенты, статистически значимые на уровнях 10% (*), 5% (**) и 1% (***). Прочерки означают, что соответствующая переменная не входила в спецификацию соответствующей модели.

В таблице 8 в строках используются следующие обозначения: «Константа» показывает значение константы в регрессии, «Импорт» – коэффициент, перед дамми переменной, являющейся индикатором импортной активности компании, он отражает уровень влияния импортной деятельности на производительность. «Экспорт», как и «Импорт», показывает значение коэффициента перед дамми переменной, являющейся индикатором экспортной деятельности. Данный коэффициент характеризует меру влияния участия в экспортной деятельности на производительность; переменная «Импорт*Экспорт» показывает значение коэффициента перед произведением индикаторов участия в экспортной и импортной деятельностей, она отражает силу нелинейного совместного взаимодействия импортной и экспортной деятельности фирмы на её уровень эффективности. «Труд» – коэффициент перед логарифмом труда фирмы. «Капитал» – коэффициент, перед логарифмом капитала фирмы.

Таблица 8 показывает результаты оценки регрессии, с дополнительными переменными, показывающими логарифм труда и капитала фирм. Добавление данных переменных не меняет качественно результаты, полученные без данных переменных. Модели, в которых добавляются нелинейные слагаемые по индикаторам экспортной и импортной деятельностей, показывают, что в них значение коэффициентов при линейных слагаемых у индикаторов уменьшается. Следовательно, она также указывает на присутствие нелинейных связей между индикаторами экспортной и импортной деятельностей. В тоже время, значения коэффициентов перед логарифмами труда и капитала меняются незначительно при добавлении в модель контролей на различные индикаторы экспортных статусов.

Таблица 8 – Значение оценки коэффициентов в регрессии эффективности фирмы на индикатор внешнеэкономической деятельности

	Модель1 (экспорт)	Модель2 (импорт)	Модель3 (экспорт* импорт)	Модель4 (экспорт + импорт)	Модель5 (экспорт + импорт + экспорт*импорт)

Константа	44.001*** (0.045)	42.231*** (0.055)	44.084*** (0.049)	45.52*** (0.066)	42.327*** (0.047)
Экспорт	7.531*** (0.114)	-	-	5.722*** (0.125)	2.325*** (0.202)
Импорт	-	7.615*** (0.093)	-	7.493*** (0.112)	4.258*** (0.196)
Импорт*Экспорт	-	-	9.005*** (0.084)	-	4.744*** (0.202)
Труд	-1.012*** (0.015)	-0.97*** (0.027)	-1.322*** (0.031)	-1.347*** (0.024)	-1.349*** (0.015)
Капитал	0.110*** (0.021)	-0.062*** (0.022)	-0.102*** (0.021)	-0.133*** (0.017)	-0.111*** (0.023)

Примечание — Результаты получены с помощью метода наименьших квадратов, с использованием предложенной в работе меры эффективности в качестве зависимой переменной. В скобках указаны стандартные ошибки оценок, устойчивые к гетероскедастичности и кластеризации ошибок. Звездочками отмечены коэффициенты, статистически значимые на уровнях 10% (*), 5% (**) и 1% (***). Прочерки означают, что соответствующая переменная не входила в спецификацию соответствующей модели.

В таблице 9 в строках используются следующие обозначения: «Константа» показывает значение константы в регрессии, «Импорт» – коэффициент, перед дамми переменной, являющейся индикатором импортной активности компании, он отражает уровень влияния импортной деятельности на производительность. «Экспорт», как и «Импорт», показывает значение коэффициента перед дамми переменной, являющейся индикатором экспортной деятельности, данный коэффициент характеризует меру влияния участия в экспортной деятельности на производительность; переменная «Импорт*Экспорт» показывает значение коэффициента перед произведением индикаторов участия в экспортной и импортной деятельностей, она отражает силу нелинейного совместного взаимодействия импортной и экспортной деятельности фирмы на её уровень эффективности. «Труд» – коэффициент перед логарифмом труда фирмы. «Капитал» – коэффициент, перед логарифмом капитала фирмы.

Таблица 9 показывает значения регрессионных коэффициентов при дополнительных контролях на регионы, к которым принадлежат фирмы. Коэффициенты перед индикаторами регионов не включены в таблицу, чтобы её не перегружать. Видно, что учёт регионов в качестве контролей не приводит к качественному изменению результатов. Коэффициенты поменялись незначительно по сравнению с начальной регрессионной

моделью. Также наблюдается понижение оценки линейного эффекта индикаторов участия в экспортной деятельности при учёте их нелинейного взаимодействия.

Таблица 9 – Значение оценки коэффициентов в регрессии эффективности фирмы на индикатор внешнеэкономической деятельности

	Модель1 (экспорт)	Модель2 (импорт)	Модель3 (экспорт* импорт)	Модель4 (экспорт + импорт)	Модель5 (экспорт + импорт + экспорт*импорт)
Константа	42.052*** (0.113)	43.311*** (0.222)	42.954*** (0.193)	44.001*** (0.218)	43.662*** (0.219)
Экспорт	8.217*** (0.99)	-	-	5.002*** (0.102)	2.753*** (0.137)
Импорт	-	9.743*** (0.076)	-	8.433*** (0.11)	4.754*** (0.201)
Импорт*Экспорт	-	-	8.927*** (0.079)	-	4.895*** (0.207)
Труд	-1.019*** (0.021)	-1.322*** (0.019)	-1.415*** (0.019)	-1.243*** (0.022)	-1.153*** (0.024)
Капитал	0.045* (0.010)	-0.124*** (0.016)	-0.157*** (0.009)	-0.184*** (0.013)	-0.152*** (0.013)

Примечание — Результаты получены с помощью метода наименьших квадратов с использованием предложенной в работе меры эффективности в качестве зависимой переменной. В скобках указаны стандартные ошибки оценок, устойчивые к гетероскедастичности и кластеризации ошибок. Звездочками отмечены коэффициенты, статистически значимые на уровнях 10% (*), 5% (**) и 1% (***). Прочерки означают, что соответствующая переменная не входила в спецификацию соответствующей модели.

В таблице 10 в строках используются следующие обозначения: «Константа» показывает значение константы в регрессии, «Импорт» – коэффициент, перед дамми переменной, являющейся индикатором импортной активности компании, он отражает уровень влияния импортной деятельности на производительность. «Экспорт», как и «Импорт», показывает значение коэффициента перед дамми переменной, являющейся индикатором экспортной деятельности, данный коэффициент характеризует меру влияния участия в экспортной деятельности на производительность; переменная «Импорт*Экспорт» показывает значение коэффициента перед произведением индикаторов участия в экспортной и импортной деятельностей, она отражает силу нелинейного совместного взаимодействия импортной и экспортной деятельности фирмы на её уровень эффективности. «Труд» –

коэффициент перед логарифмом труда фирмы. «Капитал» – коэффициент, перед логарифмом капитала фирмы.

Таблица 10 – Значение оценки коэффициентов в регрессии эффективности фирмы на индикатор внешнеэкономической деятельности, логарифм труда и капитала и год измерения

	Модель1 (экспорт)	Модель2 (импорт)	Модель3 (экспорт* импорт)	Модель4 (экспорт + импорт)	Модель5 (экспорт + импорт + экспорт*импорт)
Константа	40.221*** (0.121)	41.687*** (0.121)	42.217*** (0.121)	42.111*** (0.121)	42.005*** (0.121)
Экспорт	7.16*** (0.122)	-	-	4.521*** (0.108)	2.215*** (0.152)
Импорт	-	7.997*** (0.072)	-	7.01*** (0.072)	3.814*** (0.182)
Импорт*Экспорт	-	-	8.551*** (0.065)	-	4.326** (0.194)
2012 год	0.752*** (0.141)	0.733*** (0.141)	0.692*** (0.141)	0.734*** (0.141)	0.741*** (0.141)
2013 год	0.532*** (0.128)	0.533*** (0. 128)	0.541*** (0.128)	0.495*** (0.128)	0.506*** (0.128)
2014 год	0.603*** (0.131)	8.01*** (0.131)	8.101*** (0.131)	8.002*** (0.131)	0.792*** (0.131)
2015 год	-1.738*** (0.118)	-1.827*** (0.118)	-1.723*** (0.118)	-1.562*** (0.118)	-1.703*** (0.118)
2016 год	-2.601*** (0.128)	-2.422*** (0.128)	-2.293*** (0.128)	-2.311*** (0.128)	-2.331*** (0.128)
2017 год	6.552*** (0.123)	6.63- (0.123)	6.54- (0.123)	6.600*** (0.123)	6.522*** (0.124)
2018 год	4.723- (0.142)	4.422*** (0.142)	4.812- (0.142)	4.732*** (0.142)	4.731*** (0.142)
Труд	-0.343*** (0.031)	-0.44*** (0.031)	-0.532*** (0.030)	-0.567*** (0.031)	-0.566*** (0.030)
Капитал	-0.201 (0.012)	-0.265 (0.012)	-0.301 (0.012)	-0.331 (0.012)	-0.324 (0.012)

Примечание — Результаты получены с помощью метода наименьших квадратов, используя предложенную в работе меру эффективности в качестве зависимой переменной. В скобках указаны стандартные ошибки оценок, устойчивые к гетероскедастичности и кластеризации ошибок. Звездочками отмечены коэффициенты, статистически значимые на уровнях 10% (*), 5% (**) и 1% (***). Прочерки означают, что соответствующая переменная не входила в спецификацию соответствующей модели.

В таблице 11 в столбцах показаны отрасли, к которым относятся компании, на основе ОКВЭД и средний эффект воздействия на их уровень производительности от участия в экспортной деятельности. Чтобы сократить таблицу, она была разделена на две части и соединена по горизонтали, в каждой строке написана отрасль и значение среднего эффекта для данной отрасли.

В таблице 11 представлены результаты оценивания среднего эффекта воздействия от выхода компаний на экспортный рынки, в зависимости от отрасли, согласно ОКВЭД. Из неё видно, что наибольший прирост уровня производительности, от экспортной деятельности, получают компании, занимающиеся производством табачных изделий и одежды. Наименьшему эффекту подвержены компании, занимающиеся производством лекарственных средств и материалов, применяемых в медицинских целях и производством компьютеров, электронных и оптических изделий.

Таблица 11 – Средний эффект от выхода на экспортный рынок по отраслям

Отрасль	Средний эффект	Отрасль	Средний эффект
Производство табачных изделий	8.55	Производство прочих готовых изделий	3.31
Производство одежды	7.19	Деятельность полиграфическая и копирование носителей информации	3.3
Производство напитков	6.15	Производство мебели	3.27
Производство кокса и нефтепродуктов	5.51	Производство готовых металлических изделий, кроме машин и оборудования	3.05
Производство автотранспортных средств, прицепов и полуприцепов	5.24	Производство прочих транспортных средств и оборудования	2.58
Производство химических веществ и химических продуктов	4.25	Производство кожи и изделий из кожи	2.53

Производство пищевых продуктов	4.05	Производство электрического оборудования	2.45
Производство текстильных изделий	4.05	Производство прочей неметаллической минеральной продукции	2.45

Продолжение таблицы 15

Обработка древесины и производство изделий из дерева и пробки, кроме мебели	3.8	Производство машин и оборудования, не включенных в другие группировки	2.32
Ремонт и монтаж машин и оборудования	3.78	Производство бумаги и бумажных изделий	2.02
Производство резиновых и пластмассовых изделий	3.55	Производство лекарственных средств и материалов, применяемых в медицинских целях	2.02
Производство металлургическое	3.46	Производство компьютеров, электронных и оптических изделий	1.08

Примечание: результаты получены посредством усреднения индивидуальных эффектов воздействия по компаниям, принадлежащим соответствующим отраслям в соответствии с ОКВЭД.

ЗАКЛЮЧЕНИЕ

В рамках работы были предложен и реализован подход к построению оценок последствий выхода компаний на экспортные рынки, принимающий во внимание эффект самоотбора. Для этого были рассмотрены статьи, в которых обсуждались его механизмы, а также свидетельства влияния эффекта самоотбора на выход российских предприятий на экспортные рынки.

Также был проведен эмпирический анализ эффектов от выхода компаний на экспортный рынок с учетом эффекта самоотбора. Оценки проводились с использованием процедуры двойного выбора LASSO, которая, с одной стороны, позволила построить корректные тесты на значимость и доверительные интервалы, а с другой, позволила оценить неоднородный эффект воздействия. С помощью этой процедуры были получены оценки эффектов воздействия на уровне отдельных предприятий в зависимости от их различных характеристик, что позволило установить, что эффект действительно существенно неоднороден. Например, в зависимости от размера фирмы (фирмы с меньшим числом сотрудников при прочих равных могут ожидать больший прирост производительности от выхода на экспортный рынок), кроме того, эффект существенно различается по отраслям (в частности, для предприятий в сферах производства одежды или производства напитков выход на экспортные рынки может быть в среднем связан с существенным увеличением производительности (порядка 6-7%%), а для производителей бумаги и бумажных изделий или производителей лекарственных средств средний ожидаемый эффект оказывается существенно ниже (2%).

Полученные результаты могут быть использованы для того, чтобы прогнозировать эффективность отдельных мер по облегчению доступа Российских предприятий на экспортные рынки и позволяют формулировать приоритеты в данной области, посредством определения того, в каких отраслях меры могут быть наиболее эффективны с точки зрения роста производительности. Кроме того, примененные в работе методы демонстрируют, что неоднородные эффекты воздействия могут оказывать существенную помощь при планировании экономической политики в целом посредством оценки эффектов на уровне отдельных экономических агентов, а не только на агрегированном уровне. Эта методология становится особенно важна в современных условиях доступности больших наборов данных, которые позволяют действовать намного более эффективно.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

[1] M. A. Delgado, J. C. Farinas и S. Ruano, «Firm productivity and export markets: A non-parametric approach,» *Journal of international Economics*, т. 57, № 2, p. 397–422, 2002.

[2] J. R. Tybout, «Heterogeneity and productivity growth: Assessing the evidence,» *Industrial evolution in developing countries: Micro patterns of turnover, productivity, and market structure*, p. 43–72, 1996.

[3] B.-Y. Aw и A. R.-m. Hwang, «Productivity and the export market: A firm-level analysis,» *Journal of development economics*, т. 47, № 2, p. 313–332, 1995.

[4] A. B. Bernard, J. B. Jensen и R. Z. Lawrence, «Exporters, jobs, and wages in us manufacturing: 1976-1987,» *Microeconomics*, т. 1995, p. 67–119, 1995.

[5] A. B. Bernard и J. Wagner, «Exports and success in german manufacturing,» *Weltwirtschaftliches Archiv*, т. 133, № 1, p. 134–157, 1997.

[6] B.-Y. Aw-Roberts, X. Chen и M. J. Roberts, «Firm-level evidence on productivity differentials, turnover, and exports in taiwanese manufacturing,» *National Bureau of Economic Research Cambridge*, 1997.

[7] B. Y. Aw, S. Chung и M. J. Roberts, «Productivity and turnover in the export market: Micro-level evidence from the republic of korea and taiwan (china),» *The World Bank Economic Review*, т. 14, № 1, p. 65–90, 2000.

[8] A. A. I. L. D. O. Giammario Impullitti, «A theory of entry into and exit from export markets,» *Journal of International Economics* , pp. 75-90, 2013.

[9] M. M., «The Impact of Trade on Intra-Industry reallocation and Aggregate Industry Productivity,» *Econometrica*, pp. 1965-1725, 2003.

[10] A. Belloni, V. Chernozhukov и C. Hansen, «High-dimensional methods and inference on structural and treatment effects,» *Journal of Economic Perspectives*, т. 28, № 2, p. 29–50, 2014.

[11]]. J. H. Friedman, «The elements of statistical learning: Data mining, inference, and prediction,» *springer open*, 2017.

[12] H. Leeb и B. M. Pötscher, «developments in model selection and related areas,» *Econometric Theory*, т. 24, № 2, p. 319–322, 2008.

[13] X. Chen, «Large sample sieve estimation of semi-nonparametric models,» *Handbook of econometrics*, т. 6, p. 5549–5632, 2007.

[14]]. L. E. Frank и J. H. Friedman, «A statistical view of some chemometrics regression tools,» *Technometrics*, т. 35, № 2, p. 109–135, 1993.

[15] R. Tibshirani, «Regression shrinkage and selection via the lasso,» *Journal of the Royal Statistical Society: Series B (Methodological)*, т. 58, № 1, p. 267–288, 1996.

[16] A. Belloni, D. Chen, V. Chernozhukov и C. Hansen, «Sparse models and methods for optimal instruments with an application to eminent domain,» *Econometrica,* т. 80, № 6, p. 2369–2429, 2012.

[17] P. J. Bickel, Y. Ritov и A. B. Tsybakov, «Simultaneous analysis of lasso and dantzig selector,» *The Annals of statistics,* т. 37, № 4, p. 1705–1732, 2009.

[18] S. Ng и J. Bai, «Selecting instrumental variables in a data rich environment,» *Journal of Time Series Econometrics,* т. 1, № 1, 2009.

[19] P. M. Robinson, «Root-n-consistent semiparametric regression,» *Econometrica: Journal of the Econometric Society,* p. 931–954, 1988.

[20] J. J. D. III и S. D. Levitt, «The impact of legalized abortion on crime,» *The Quarterly Journal of Economics,* т. 116, № 2, p. 379–420, 2001.

[21] D. Acemoglu, S. Johnson и J. A. Robinson, «The colonial origins of comparative development: An empirical investigation,» *American economic review,* т. 91, № 5, p. 1369–1401, 2001.

[22] A. Belloni, V. Chernozhukov и C. Hansen, «Inference on treatment effects after selection amongst high-dimensional controls,» *The Review of Economic Studies,* p. 608–650, 2014.

[23] V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey и J. Robins, «Double/debiased machine learning for treatment and structural parameters.,» *The Econometrics Journal,* т. 21, № 1, pp. 1-68, 2018.

[24] S. Athey и G. Imbens, «Recursive partitioning for heterogeneous causal effects.,» *Proceedings of the National Academy of Sciences,* т. 113, № 27, pp. 7353-7360, 2016.

[25] S. (Wager и S. Athey, «Estimation and inference of heterogeneous treatment effects using random forests.,» *Journal of the American Statistical Association,* т. 113, № 523, pp. 1228-1242, 2018.

[26] Н. А. Краснопеева, Е. Ю. Назруллаева, А. А. Пересецкий и Е. И. Щетинин, «. "Экспортировать или нет? Экспортный статус и техническая эффективность российских предприятий.,» *Вопросы экономики,* т. 7, pp. 123-146, 2016.

[27] D. Aigner, C. K. Lovell и P. Schmidt, «Formulation and estimation of stochastic frontier production function models,» *{Journal of econometrics,* т. 6, № 1, pp. 21--37, 1977.

[28] W. Meeusen и J. v. D. Broeck, «Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error,» *International Economic Review,* т. 18, № 2, pp. 435--444, 1977.

[29] A. B. Bernard и J. B. Jensen, «Exceptional exporter performance: cause, effect, or both?,» *Journal of international economics,* № 47.1, pp. 1-25, 1999.

[30] L. Halpern, M. Koren и A. Szeidl, «Imported inputs and productivity,» *American Economic Review,* т. 105, № 12, pp. 3660--3703, 2015.

[31] F. Caselli и D. J. Wilson, «Importing technology,» *Journal of monetary Economics,* т. 51, № 1, pp. 1--32, 2004.

[32] I. Bertschek, J. Hogrefe и F. Rasel, «Trade and technology: new evidence on the productivity sorting of firms,» *Review of World Economics,* т. 151, № 1, pp. 53--72, 2015.

[33] H. Leeb и B. Pötscher, «Model Selection And Inference: Facts And Fiction,» *Econometric Theory,* pp. 21-59, 2005.

[34] Krugman и P. R., «Increasing returns, monopolistic competition, and internatioanal trade.,» *Economics,* pp. 469-479, 1979.

[35] H. A. и Hpenhayn, «Entry, exit, and firm dynamics in long run equilibrium,» *Econometrica: Journal of the Econometric Society,* pp. 1127-1150, 1992.

[36] A. K., Dixit, J. E. и Stiglitz, «Monopolistic competition and optimum product diversity,» *The American economic review,* p. 297—308, 1977.

[37] C. Pinho и L. Martins, «Exporting barriers: Insights from portuguese small-and medium-sized exporters and non-exporters,» *Journal of international Entrepreneurship,* т. 8, № 3, p. 2010, 2010.

[38] S. Rabino, «An examination of barriers to exporting encountered by small manufacturing companies,» *Management International Review,* p. 67–73, 1980.

[39] R. C. Hook и M. R. Czinkota, «Export activities and prospects of hawaiian firms,» *International Marketing Review,* 1988.

[40] T. W. Sharkey, J.-S. Lim и K. I. Kim, «Export development and perceived export barriers: An empirical analysis of small firms,» *Management International Review,* p. 33–40, 1989.

[41]]. L. C. Leonidou, «An analysis of the barriers hindering small business export development,» *Journal of small business management,* т. 42, № 3, p. 279–302, 2004.

[42] G. Tesfom и C. Lutz, «A classification of export marketing problems of small and medium sized manufacturing firms in developing countries,» *International Journal of Emerging Markets,* 2006.

[43] C. Pinho и L. Martins, «Exporting barriers: Insights from portuguese small-and medium-sized exporters and non-exporters,» *Journal of international Entrepreneurship,* т. 8, № 3, p. 254–272, 2010.

[44] K. J. Ruhl и J. L. Willis, «New exporter dynamics,» *International Economic Review,* т. 58, № 3, p. 703–726, 2017.

[45] J. Eaton, M. Eslava, C. Krizan, D. Jinkins и J. T. J., «A search and learning model of export dynamics, kunpublished,» *Pennsylvania State University,* 2014.

[46] Y. Rho и J. Rodrigue, «Firm-level investment and export dynamics,» *International Economic Review,* т. 57, № 1, p. 271–304, 2016.

[47] D. Kohn, F. Leibovici и M. Szkup, «Financial frictions and new exporter dynamics,» *International economic review,* т. 57, № 2, p. 453–486, 2016.

[48] L. Foster, J. Haltiwanger и C. Syverson, «The slow growth of new plants: Learning about demand?,» *Economica,* т. 83, № 329, p. 91–129, 2016.

[49] A. B. Bernard, J. Eaton, J. B. Jensen и S. Kortum, «Plants and productivity in international trade,» *American economic review,* т. 93, № 4, p. 1268–1290, 2003.

[50] A. Irarrazabal, A. Moxnes и L. D. Opromolla, «The margins of multinational production and the role of intrafirm trade,» *Journal of Political Economy,* т. 121, № 1, p. 74–126, 2013.

[51] Das, Sanghamitra, Roberts, M. J., Tybout и J. R, «Market entry costs, producer heterogeneity, and export dynamics,» *Econometric,* т. 75, № 3, pp. 837--873, 2007.

[52] Eaton, S. Kortum и F. Kramarz, «An anatomy of international trade: Evidence from french firms,» *Econometrica,* т. 79, № 5, p. 1453–1498, 2011.

[53] P. K. R. Baldwin, «Persistent trade effects of large exchange rate shocks,» *The Quarterly Journal of Economics,* т. 104, № 4, p. 635–654, 1989.

[54] M. J. Roberts и J. R. Tybout, «The decision to export in colombia: An empirical model of entry with sunk costs,» *The American Economic Review,* p. 545–564, 1997.

[55] B. Alexandre, ChernozhukovVictor и H. Christian, «Inference on Treatment Effectsafter Selection among High-Dimensional Controls,» *Review of Economic Studies,* т. 81, p. 608–650, 2014.

[56] А. А. Федюнина и Ю. В. Аверьянова, «Эмпирический анализ факторов конкурентоспособности российских экспортеров,» *Экономическая политика,* т. 13, № 6, 2018.

[57] S. Poupakis, «Does FDI in upstream and downstream sectors facilitate quality upgrading? Evidence from russian exporters,» *Oxford Bulletin of Economics and Statistics,* 2021.