

Влияние искусственного интеллекта на российский рынок труда

Скоробогатов А.С., Бондаренко Д.И., Свиридов О.И.

Последние десятилетия мы наблюдаем стремительное развитие в области искусственного интеллекта (ИИ). И в последние годы по большей мере это связано с развитием методов машинного обучения и доступности огромных массивов данных. Искусственный интеллект, или машинное обучение, относится к алгоритмам, которые учатся выполнять задачи, выявляя статистические закономерности в данных и не следуя инструкциям, предоставляемым людьми. Эта технология недавно достигла сверхчеловеческой производительности в широком спектре экономически ценных задач. И справедливо ожидать, что данный рост только усилится [7], но при этом людям до сих пор не известны его масштабы и конечный результат. И все же, даже самый негативный итог с подчинением человека сверхинтеллекту можно избежать. Предполагается [16], что в этом случае мы можем создать ИИ на новых принципах, основным из которых будет неоднозначность по человеческим предпочтениям, которым ИИ должен соответствовать. В таком случае, ИИ был бы скромным, альтруистичным и направленным на достижение целей человека, а не своих собственных. Возвращаясь в настоящее, справедливо ожидать, что алгоритмы ИИ¹, как и любая новая технология, трансформирует рынок труда. А значит, можно ожидать изменений относительных требований к профессиям [8], повышения производительности и общего уровня жизни. Понимание этих последствий важно для многих целей. Например, это позволяет правительственным органам разрабатывать соответствующую политику в области образования и

¹ Строго говоря, в статье речь идет скорее об отдельной категории алгоритмов машинного обучения. Тем не менее мы решили следовать терминологии, используемой в уже опубликованных работах, и называть это «алгоритмы ИИ»

рынка труда, а также помогает отдельным людям делать правильный выбор карьерного пути.

Логично будет предположить, что внедрение подобных технологий в первую очередь поставит под угрозу существования профессии, основанные на циклических процедурах и алгоритмах. Это верно и уже можно заметить для таких профессий как: продавец, водитель или машинист (в разных городах мира уже сейчас имеется беспилотное метро), курьер и работник складской логистики, сотрудник банка, принимающий заявки на кредиты, оператор колл-центра. Так же алгоритмы ИИ активно внедряют в медицине [14] для выявления диагнозов, подбора лечения пациентов и разработки лекарств. Влияние заметно и в сфере бизнеса [15], и финансов. Яркий пример: Сбербанк. Подавляющим число заявок на кредиты занимается алгоритм ИИ. Изучаются кредитные истории клиентов, уровни дохода и затрат. Кроме преимуществ в скорости выполнения и точности в сравнении с человеком, так же можно отметить процедуру онлайн-финансирования. Опираясь на информацию по транзакциям, историю кредитования клиента и другие данные, ИИ вычисляет, когда у клиента возникает потребность в финансировании, и заранее предлагает ему необходимую сумму. Но при всем этом, в компании в данный момент не стоит вопрос полного замещения человека алгоритмами ИИ, поскольку алгоритм ИИ не сможет сделать корректный прогноз в отношении компаний, которые недавно начали свою деятельность и бессилён при прогнозе будущего проектов компаний, которые прошли процедуру реорганизации или слияния.

Данная трансформация рынка труда может обеспечить как будущее без работы [18], так и повышение производительности труда и опыта работы людей [10]. Таким образом, в текущий момент, согласно имеющимся исследованиям, можно заявить, что существующие данные не являются исчерпывающими. И все еще нету систематических доказательств того, действительно ли произошло значительное увеличение внедрения алгоритмов

ИИ, о котором часто утверждают средства массовой информации. Ведь можно найти примеры того, как технологии ИИ либо заменяют рабочих, либо дополняют их, ведь алгоритмы ИИ, как технологическая платформа, способны делать и то, и другое. Тем не менее, скорость внедрения алгоритмов ИИ неравномерна, и число рабочих мест, на которые повлиял ИИ, зависит от уровня наличия высокотехнологичных производств и филиалов крупных компаний в конкретном месте [5].

Исследования, посвященные влиянию именно алгоритмов ИИ, а не автоматизации в целом, начали появляться относительно недавно и их число постоянно растет наравне с исследованиями остальных областей автоматизации. В данной работе изучается внедрение алгоритмов искусственного интеллекта в России их влияние на рынок труда. Идея состоит в том, что внедрение ИИ можно частично классифицировать по следам, которые оно оставляет в фирмах, рассматривая найм работников, специализирующихся на деятельности, связанной с алгоритмами ИИ, например, машинный перевод или распознавание изображений. Для реализации этой идеи был создан набор данных об активности ИИ на уровне фирмы, основанный на онлайн-объявлениях о вакансиях в HeadHunter, содержащих подробные требования к навыкам по вакансиям.

Рассмотрение данных на уровне фирмы также полезно для понимания того, как автоматизация в целом влияет на производственный процесс и производительность [11]. Согласно данным по Франции, внедрение роботов на уровне фирм совпадает со снижением доли рабочей силы, увеличением добавленной стоимости и производительности, а также снижением доли производственных рабочих [2]. Общая занятость растет быстрее в фирмах, внедряющих роботов. Этот положительный эффект может быть следствием перераспределения производства и рабочей силы в пользу фирм, которые снижают свои издержки по сравнению со своими конкурентами. Данное

перераспределение объясняет положительное влияние роботов на уровне фирмы.

Таким образом, автоматизация заменяет капиталом задачи, ранее выполнявшиеся трудом, уменьшая долю труда в добавленной стоимости и увеличивая добавленную стоимость на одного работника в процессе. В то время как более высокая производительность за счет автоматизации, как правило, увеличивает спрос на рабочую силу, ее эффект вытеснения может перевесить это положительное влияние и может привести к общему снижению занятости и заработной платы [3].

Наша работа довольно во многом перекликается с другой работой Acemoglu et al. [1], в которой рассматривается воздействие уже именно алгоритмов ИИ на уровне фирм в США с 2010 по 2018 годы. В ней аналогично утверждается, что фирмы, структура задач которых позволяет использовать алгоритмы ИИ, существенно увеличили количество вакансий, связанных с ними. Так же получено, что воздействие алгоритмов ИИ связано с значительными изменениями в наборе навыков профессий, по соответствующим вакансиям, и что внедрение ИИ оказывает реальное влияние на количество рабочих мест, не связанных с ИИ, позволяя фирмам заменять некоторые задачи, ранее выполнявшиеся работниками, делая определенные навыки излишними, одновременно создавая спрос на новые навыки. Интерес представляет проверка указанных выводов для Российских данных.

В теоретическом разделе описано расширение модели, используемой в статье Acemoglu et al. [1], различающее фирмы по индустриям. Оно состоит в том, что одни и те же алгоритмы ИИ будут иметь разную производительность в различных индустриях. Предположение о зависимости производительности алгоритмов от индустрии, в которую они внедряются, является достаточно интуитивным.

Во-первых, это объясняется различными масштабами применения. Чем больше масштабы применения ИИ, тем больше требований к данным и производительности железа будут предъявлять алгоритмы и тем тяжелее будет повышать их производительность. Так, например, чат-бот, созданный для магазина или сети магазинов, будет при прочих равных иметь бóльшую потенциальную производительность по сравнению с чат-ботом для крупного банка. Причиной этого является именно масштаб применения: в банке, с его разнообразием и сложностью предоставляемых услуг, сложность и разнообразие запросов, направляемых клиентами чат-боту, будет гораздо больше.

Во-вторых, применение одной и той же технологии в разных индустриях может отличаться по своей специфике. Например, технологию биометрического ключа применяют как производители смартфонов, так и банки и оборонно-промышленные комплексы. Однако совершенно очевидно, что в первом случае усилия разработчиков направлены на максимальное упрощение (удешевление) алгоритма, что сопровождается понижением его точности. В то же время разработка технологии биометрического ключа в банках или военных комплексах фокусируется исключительно на безопасности, т.е. на точность. Даже, если и в том, и в другом случае применяется технология нейронной сети, то, как минимум, гиперпараметры у таких алгоритмов будут различаться, а значит и производительность может быть разной.

В-третьих, различные стохастические элементы в процессе производства. Некоторые индустрии сталкиваются с шоками производства гораздо реже, чем другие, что позволяет им собирать менее зашумленные данные, что также неизбежно скажется на работе многих алгоритмов.

В дальнейшем статья организована следующим образом: в разделе 1 представлена теоретическая модель, мотивирующая нашу эмпирическую стратегию. В разделе 2 описываются данные, а в разделе 3 представлена уже

сама эмпирическая стратегия. В разделе 4 представлены основные результаты по воздействию ИИ на найм на уровне фирм. В последнем разделе представлены выводы и заключения. Эмпирические результаты и графики представлены в Приложении.

Базовая модель

Мы используем теоретическую модель, аналогичную работе Acemoglu et al. [1] для описания гипотез и интерпретации эмпирических результатов. В данном разделе будет представлена краткая выжимка со всеми необходимыми уравнениями. Более подробное описание представлено в Acemoglu et al. [1]

Основная суть модели сводится к разделению выпуска y_f каждой фирмы f на процессы решения определенных задач $y_f(x)$, где x – конкретная задача из множества всех существующих задач T . Производственная функция каждой фирмы описывается уравнением Кобба-Дугласа:

$$\ln(y_f) = \int_{T_f} \alpha(x) \ln(y_f(x)) dx ; \quad (1)$$

где T – множество всевозможных задач, T_f – множество всевозможных задач, реализуемых фирмой f в процессе производства, $\alpha(x) \geq 0$ – коэффициент «важности» задачи x в производственном процессе фирмы, при чем функция имеет постоянную отдачу от масштаба, т.е. $\int_{T_f} \alpha(x) dx = 1$ для любого T_f .

Сам же процесс решения задач осуществляется за счет использования двух факторов производства: человеческого труда $l_f(x)$ и алгоритмов искусственного интеллекта (ИИ) $a_f(x)$. Функция реализации задач x описывается следующим уравнением:

$$y_f(x) = A_f \left[(\varphi_l(x) l_f(x))^{\frac{\varepsilon-1}{\varepsilon}} + (\varphi_a(x) a_f(x))^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}} ; \quad (2)$$

где A_f – уровень технологического развития фирмы f , ε – эластичность замещения между человеческим трудом и алгоритмами, $\varphi_l(x), \varphi_a(x)$ – коэффициенты, отражающие качество (скорость, точность и т.д.) выполнения задания x человеческим капиталом и алгоритмами соответственно и предполагаются постоянными для всех фирм. Человеческий капитал может

выполнить любую задачу из множества T , а алгоритмы способны выполнять только часть задач $T^a \in T$.

Мы предполагаем, что множество заданий T разделено на подмножества задач, которые выполняются конкретными профессиями. Иными словами, существует множество профессий O такое, что каждая профессия $o \in O$ имеет свой определенный набор задач $T^o \in T$, который данная профессия способна выполнять. Таким образом структура задач, выполняемых в производственном процессе фирмы f , будет определяться структурой профессий работников, нанятых в этой фирме: $T_f = \cup_{o \in O_f} T^o$.

Данные предположения делают спрос на факторы производства зависимым только от структуры выполняемых в фирме задач. То есть если предположить, что качество выполнения алгоритмами определенного множества задач выросло, то изменения спроса на факторы производства среди фирм будет гетерогенным и будет однозначно определяться долей профессий, выполняющих эти задачи в производстве фирмы. Иными словами, прогресс алгоритмов ИИ будет влиять больше на те фирмы, в производственном процессе которых большая доля задач из множества T^a . Тогда индекс влияния алгоритмов ИИ на фирму можно описать как долю таких задач в множестве всех задач фирмы:

$$ExpAI = \frac{\int_{x \in T_f \cap T^a} l_f(x) dx}{\int_{x \in T_f} l_f(x) dx}. \quad (3)$$

Основная проблема такого индекса заключается в том, что его достаточно сложно оценить явным образом. Мы используем индекс влияния алгоритмов ИИ на профессии, посчитанный в работе [17] чтобы оценить $ExpAI$. В обозначенных выше терминах посчитанный индекс будет описываться следующей формулой:

$$OExp^o = \frac{\int_{x \in T^o \cap T^a} l(x) dx}{\int_{x \in T^o} l(x) dx}; \quad (4)$$

где $l(x)$ – средняя занятость по задаче x . Иными словами, индекс влияния ИИ на профессию обозначает долю занятости профессии по тем задачам, которые может выполнять ИИ, в общей занятости в профессии. Чем больше данный индекс, тем, соответственно больше занятого человеческого капитала в данной профессии можно заменить на работающий алгоритм.

Обозначим среднюю занятость в профессии как $l^o = \int_{x \in T^o} l(x) dx$. Из предположения об однородности фирм (общая технология производства) следует, что $l(x) = l_f$. Тогда можем записать:

$$\frac{\sum_{o \in O_f} OExp^o l^o}{\sum_{o \in O_f} l^o} = \frac{\sum_{o \in O_f} \frac{\int_{x \in T^o \cap T^a} l(x) dx}{\int_{x \in T^o} l(x) dx} l^o}{\sum_{o \in O_f} l^o} = \frac{\int_{x \in T_f \cap T^a} l_f(x) dx}{\int_{x \in T_f} l_f(x) dx} = ExpAI_f. \quad (5)$$

То есть индекс влияния алгоритмов ИИ на фирму можно представить как сумму индексов профессий, взвешенных на долю данных профессий в общей занятости фирмы. Эта формула будет использована нами для оценки $ExpAI_f$.

Расширение модели

Предположение об однородности фирм является оптимальным с точки зрения демонстрации общей связи между индексом влияния ИИ на фирмы и индексом влияния ИИ на профессии. Тем не менее, оно может показаться слишком сильным и далеким от жизни. Мы ослабим положение об однородности фирм, предположив, что фирмы отличаются друг от друга по индустриям. Отличие будет выражаться в разной производительности алгоритмов по одним и тем же заданиям и разной ценой внедрения алгоритмов в разные индустрии. Труд, как и раньше, предполагается однородным. Внутри индустрий фирмы однородны. Считая, что фирмы максимизируют производство каждой задачи при ограниченном бюджете, получаем следующую задачу оптимизации:

$$y_{f,i}(x) = A_{f,i} \left[\left(\varphi_l(x) l_{f,i}(x) \right)^{\frac{\varepsilon-1}{\varepsilon}} + \left(\varphi_{a,i}(x) a_{f,i}(x) \right)^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}} \xrightarrow{l,a} \max(6)$$

$$s. t. w l_{f,i}(x) + p_{a,i} a_{f,i}(x) = Y_{f,i}(x); (7)$$

где индекс i обозначает индустрию, к которой принадлежит компания f ; w, p_i – цены человеческого капитала для задачи x и цена внедрения алгоритма в индустрии i для задачи x соответственно; $Y_{f,i}(x)$ – бюджет, выделяемый фирмой для выполнения задачи x . Тогда оптимальный спрос на труд по задаче x будет:

$$l_{f,i}^*(x) = \frac{Y_{f,i}(x) \varphi_{a,i}(x)}{w \varphi_{a,i}(x) + p_{a,i}^{1-\varepsilon} \varphi_l(x) w^\varepsilon}. (8)$$

Это уравнение позволяет сделать некоторые предположения о связи между занятостью на рынке труда и производительностью алгоритмов, которые мы обсудим в следующем разделе.

В терминах данного расширения средняя занятость $l(x)$ из уравнения (4) можно представить как:

$$l(x) = \frac{1}{I} \sum_i l_i(x), (9)$$

где $l_i(x)$ – средняя занятость по задаче x в индустрии i ; I – количество индустрий в экономике. Тогда средняя занятость в профессии будет:

$$l^o = \int_{x \in T^o} l(x) dx = \int_{x \in T^o} \frac{1}{I} \sum_i l_i(x) dx = \frac{1}{I} \sum_i \int_{x \in T^o} l_i(x) dx = \frac{1}{I} \sum_i l_i^o; (10)$$

где l_i^o – средняя занятость по профессии o в индустрии i .

По аналогии с уравнением (4) покажем индекс влияния алгоритмов ИИ на профессию внутри одной индустрии:

$$OExp_i^o = \frac{\int_{x \in T^o \cap \Gamma^a} l_i(x) dx}{\int_{x \in T^o} l_i(x) dx}. (11)$$

В то же время общий индекс влияния ИИ на профессию o будет:

$$OExp^o = \frac{\int_{x \in T^o \cap T^a} l(x) dx}{\int_{x \in T^o} l(x) dx} = \frac{\int_{x \in T^o \cap T^a} \frac{1}{I} \sum_i l_i(x) dx}{\int_{x \in T^o} \frac{1}{I} \sum_i l_i(x) dx}. \quad (12)$$

Из уравнения (5) следует, что:

$$\frac{\sum_{o \in O_f} OExp_i^o l_i^o}{\sum_{o \in O_f} l_i^o} = \frac{\int_{x \in T_f \cap T^a} l_f(x) dx}{\int_{x \in T_f} l_f(x) dx} = ExpAI_f. \quad (13)$$

В то же время из уравнений (11) и (12) следует:

$$\frac{1}{I} \sum_i \frac{l_i^o}{l^o} OExp_i^o = \frac{\frac{1}{I} \sum_i \int_{x \in T^o \cap T^a} l_i(x)}{l^o} = \frac{\int_{x \in T^o \cap T^a} \frac{1}{I} \sum_i l_i(x) dx}{\int_{x \in T^o} \frac{1}{I} \sum_i l_i(x) dx} = OExp^o. \quad (14)$$

Уравнения (13) и (14) показывают, что на имеющихся данных невозможно аналогично базовой модели оценить показатели $ExpAI_f$, поскольку имеющийся у нас $OExp^o$ в данной модификации модели будет лишь агрегатором более точного показателя $OExp_i^o$, который отсутствует в данных и может оказаться сложным для сбора. По этой причине эмпирическая оценка будет проводиться на основе базовой модели.

Гипотезы

Опираясь на вышеописанную теорию, мы можем выдвинуть следующие гипотезы:

1. Фирмы с более высоким индексом влияния ИИ будут нанимать больше человеческого труда.
2. Фирмы с более высоким индексом влияния ИИ будут иметь бóльшие доли ИИ-вакансий.

Эти две гипотезы отражают противоположные по влиянию эффекты от потенциального внедрения алгоритмов в производственный процесс. Первая гипотеза является прямым следствием уравнения (8): с увеличением

производительности алгоритмов (параметра φ_a) фирмы смогут сократить свои расходы на технологичный фактор производства и перенаправить излишек на дополнительные единицы человеческого труда. Этому же уравнению можно дать и другую интерпретацию: увеличение доли алгоритмов в производственном процессе создает новые рабочие места, так как возникает потребность в человеческих ресурсах, способных поддерживать работу этих алгоритмов. Более того, предположив, что $\varepsilon < 1$ или, что $\alpha(x)$ выше у задач, которые алгоритмы способны выполнять, мы тем самым определим, что факторы являются дополняющими друг друга и алгоритмы повышают производительность рабочих. Отсюда прямо следует гипотеза 1. Иными словами, данная гипотеза говорит об эффекте дополнения человеческого капитала и алгоритмов как факторов производства.

Гипотезу 2 позволяет выдвинуть уравнение (5). Поскольку индекс влияния ИИ на фирму является по определению долей занятого труда, который может быть заменен алгоритмами, то естественным образом можем заключить, что чем больше индекс, тем больше фирма сможет сократить использование не связанного с алгоритмами человеческого труда в своем производстве, что неизбежно сократит его долю среди факторов. К таким же выводам может привести и немного иной подход: a_f рассматривается в модели сугубо как алгоритмы ИИ, но можно немного расширить понятие этого фактора и включить в него не только ИИ, но и рабочих, создающих и обслуживающих эти алгоритмы. Тогда более высокий индекс ИИ можно будет интерпретировать как возможность заменить большее количество «нетехнологичного» труда единицей «технологичного», из чего и вытекает вторая гипотеза, отражающая наличие эффекта замещения факторов друг друга.

Описанные выше эффекты являются достаточно интуитивными и находят подтверждение своего существования в реальном мире. Главный вопрос состоит в том, какой из них окажется сильнее. Существующая литература

[1],[3],[6] приходит к выводу, что таковым на данный момент оказывается эффект замещения, поскольку не наблюдается повышение найма у фирм с бóльшим индексом влияния ИИ. По этой причине гипотеза 2 в нашей работе будет основной.

Данные

В качестве источника данных был выбран HeadHunter и использовано API, позволяющее выгрузить все активные вакансии. Однако, есть одно значительное неудобство, поскольку за один запрос к серверу можно выгрузить не более двух тысяч вакансий, иначе выгрузка в целом не состоится. Данное обстоятельство значительно осложняет формирование итогового датасета. Итоговые данные были выгружены с помощью фильтров вакансий, не превышавших порогового значения в две тысячи. Основным фильтром выступало наименование индустрии. Но, поскольку одна и та же вакансия может находиться в разных индустриях, было дополнительно выгружены уникальные id вакансий, позволившие удалять дубликаты. В итоговом рабочем варианте находилось около трех с половиной миллионов вакансий для 35 тысяч фирм, чуть более сотни профессий и трёхсот индустрий.

Поскольку мы отталкиваемся от набора рабочих задач, то, чтобы подчеркнуть, что алгоритмы ИИ связаны с некоторыми конкретными из этих задач, и что фирмы, занимающиеся этими задачами, будут теми, кто использует технологии алгоритмов ИИ, нужно определить задачи, совместимые с текущими возможностями ИИ. Для этого мы рассмотрели три различных показателя: Показатель профессионального воздействия искусственного интеллекта [12][13]; Индекс пригодности для машинного обучения (SML). [9]; и оценка воздействия искусственного интеллекта Уэбба [17].

Первый из них, показатель Фелтена [13] исходил из релевантности 9 областей, в которых алгоритмы ИИ достигли значительный прогресс с 2010 года относительно 52 навыков O*NET (которые входят в каждую профессию этой базы данных с определенными коэффициентами важности и необходимого уровня владения). Для SML [9] использовалась собственная система пригодности задачи для алгоритмов ИИ состоящая из 23 критериев и примененная к базе данных профессий O*NET.

И последняя мера, Уэбба [17]. В своей работе Уэбб определяет профессии, которые подвергаются воздействию ИИ, на основе того, насколько задачи профессии, все с теми же конфидентами из O*NET, совпадают с возможностями, описанными в патентах с ИИ. Профессии, на которые приходится большая доля таких перекрывающихся задач, классифицируются как более подверженные риску.

Сравнивая данные коэффициенты, мы пришли к выводу, что индекс Уэбба дает более стабильные показатели в отношении доли вакансий и аналогично стабилен в отношении общего их числа относительно SML и меры Фелтена. Так же индекс Уэбба аналогичен индексу Фелтена по значимости и надежности, и они оба опережают в этом SML. Вместе с тем, индекс Уэбба утверждает о очень малой пригодности ИИ для административной сферы и в сфере продаж. Отчасти поэтому, наряду со своим индексом искусственного интеллекта Уэбб создал отдельный индекс воздействия программного обеспечения, относящийся к традиционному программному обеспечению, не связанному с искусственным интеллектом, которое определяет значительную пригодность программного обеспечения для офисных, административных и торговых профессий. И еще один индекс конкретно для роботов. Но, поскольку в нашем датасете, данные по продажам сильно диверсифицированы по индустриям и имеют в них очень малое количество вакансий, мы их не рассматриваем в нашей регрессии. Соответственно этот недостаток меры Уэбба не будет прослеживаться в нашем случае.

Используя этот индекс, мы строим воздействие ИИ на фирму на основе ее профессиональной структуры точно по формуле (5), рассматривая количество вакансий каждой из 35 тысяч фирм по каждой профессии.

Поскольку наша цель - изучить влияние ИИ на фирмы, использующие ИИ. Мы исключаем из рассмотрения информационно-технологическую индустрию, которая является основными поставщиками услуг ИИ, а также уже упомянутую сферу продаж.

Каждой фирме была присвоена та индустрия, которая указана в первой найденной вакансии этой фирмы. Аналогично было сделано и с городами. Количество вакансий алгоритмов ИИ по фирмам считалось аналогично Acemoglu et. al [1], по соответствию профессий с обязанностями, наиболее связанными с работой алгоритмов.

В приложениях 12-13 представлены гистограммы индекса влияния алгоритмов ИИ на фирму и его логарифма. Распределения относительно близки к нормальным. Так же построена диаграмма рассеяния числа вакансий алгоритмов ИИ по индексу влияния ИИ на фирмы, где каждой точке соответствуют примерно 1500 фирм (приложение 14). Взаимосвязь довольно близка к линейной по всему распределению и, по-видимому, не обусловлена выбросами.

Регрессионная модель

Наша эмпирическая модель описывает связь между количеством опубликованных вакансий и индексом влияния ИИ на фирму. Будет оценена следующая регрессия:

$$\ln(y_f) = \beta_0 + \beta_1 \ln(ExpAI_f) + \sum_i \gamma_i x_{f,i} + \varepsilon_f; \quad (15)$$

где y_f – количество определенных вакансий, опубликованных фирмой; $ExpAI_f$ – индекс влияния ИИ на фирму, оцененный как сумма индексов профессий, опубликованных данной фирмой, взвешенных на долю вакансий данной профессии у этой фирмы; $x_{f,i}$ – контрольные переменные на индустрию, город и размер фирмы; ε_f – переменная ошибок, включающая в себя все неучтенные факторы. Нас интересует переменная β_1 , отражающая процентное изменение объясняемой переменной при изменении индекса влияния ИИ фирмы на один процент. Мы рассмотрим 5 вариантов переменной y_f – общее количество вакансий; количество вакансий, связанных с ИИ (AI-вакансии); количество вакансий, не связанных с ИИ (Non-AI вакансии); доли AI и Non-AI вакансий. Мы разделяли вакансии на AI и Non-AI, следуя примеру Acemoglu, D. et al. [1], взяв профессии с обязанностями, наиболее связанными с работой алгоритмов.

Результаты

Оценив регрессии, мы наблюдаем положительную связь между индексом влияния ИИ и количеством опубликованных вакансий. Это тенденция характерна для всех видов u_f . В Приложениях 1-5 представлены аутпуты регрессий. Увеличение индекса влияния ИИ на фирму на 1% в среднем сопровождается увеличением общего количества вакансий фирмы на 24,7%, увеличением общего количества AI вакансий на 26,6% и Non-AI вакансий на 18,3%. В тоже время с увеличением $ExpAI_f$ на 1 % доля AI-вакансий также в среднем увеличивается на 3,4% в то время, как доля Non-AI вакансий снижается 2,5%.

Таблица 1. Сводные результаты измерений для основного датасета

	Growth of firm vacancies				
	Total vacancies	AI vacancies	Non AI vacancies	Share of AI vacancies	Share of non AI vacancies
<i>Panel A: Main data set (34 633 observations)</i>					
Firm AI exposure	0.247*** (0.013)	0.266*** (0.010)	0.183*** (0.012)	0.034*** (0.001)	-0.024*** (0.001)
Fixed Effects:					
Industry	✓	✓	✓	✓	✓
Firm location	✓	✓	✓	✓	✓
Firm size				✓	✓

Обсуждение

На первый взгляд результаты согласуются с полученными в более ранних работах: подтверждается основная гипотеза о том, что доля AI-вакансий фирм возрастает с увеличением $ExpAI_f$. Однако данные результаты говорят лишь о том, что в среднем бóльшая занятость свойственна фирмам с бóльшим индексом влияния ИИ. Но нет никаких оснований полагать, что именно потенциальное влияние алгоритмов является причинным фактором в этой закономерности. Более того, увеличение Non-AI занятости идет вразрез с существующей литературой и интуицией модели, и не дает однозначно интерпретировать результаты. Можно предположить, что в российской

экономике наблюдается сильный эффект дополнения производственных факторов и подтверждается гипотеза 1, либо что полученная значимая связь вообще является случайностью и не отражает действительность. Мы попытаемся провести ряд дополнительных проверок, которые позволят разобраться с возникшей проблемой.

Увеличение количества Non-AI вакансий с ростом индекса влияния может быть вызвано неоднородностью в данных: у нас слишком много наблюдений с нулевыми значениями количества AI-вакансий. Это, учитывая загадочный положительный знак в регрессии с Non-AI вакансиями, может объясняться тем, что фирмы с нулевыми значениями AI вакансий в большинстве своем просто не активны (в данном периоде или в принципе), и их суммарная доля в общем числе всех вакансий невелика. В итоге получается, что мы имеем относительно меньшую группу фирм (1/4), сильно выбивающуюся на фоне остальных. Имеет смысл посмотреть именно на эту группу². В приложении 6-9 видим, что результаты полностью согласуются с литературой и теорией: внутри этой группы фирм количество Non-AI вакансий никак не связано с индексом влияния ИИ, в то время как количество AI-вакансий в среднем с увеличением $ExpAI_f$ увеличивается, как и их доля в общем числе рабочих мест. Общие результаты приложений 6-9 представлены в приложении 10.

Мы объяснили неоднородность в данных повышенной деловой активностью некоей меньшей доли фирм. Нам бы хотелось усилить наше утверждение опровержением некоторых альтернативных объяснений: гетерогенность по размерам фирм и по индустриям. Следуя примеру Acemoglu et. al [1], мы взяли в качестве прокси для размера фирмы переменную общего количества вакансий. Безусловно, на кросс-секции такая прокси будет гораздо слабее, но тут мы сталкиваемся с ограничением в данных. Индустрии, итак, учтены в нашей модели и добавлена дамми переменная на каждую индустрию. В

² Выделим ее по критерию ненулевой доли AI вакансий – поиск непрофильных высококвалифицированных рабочих такого типа является неплохой прокси для активности фирмы. В качестве дополнительного аргумента отметим, что на 8000 из 34000 таких фирм приходится до 75% всех вакансий.

качестве наглядного подтверждения в приложении 11 представлен график, показывающий среднее значение $ExpAI_f$ по индустрии и количество вакансий в этой индустрии. Нетрудно увидеть, что связи там нет.

Заключение

Подводя итоги, мы получили подтверждение гипотезы 2 и не нашли подтверждения первой. Доля AI-вакансий в среднем выше в фирмах с бóльшим значением индекса влияния ИИ. В то же время, как и в более ранней литературе, у фирм с более высоким $ExrAI_f$ повышен уровень найма. Данная работа безусловно имеет ограничения, связанные с использованными в ней данными. Тем не менее, такие результаты создают подспорье для более глубокого изучения влияния алгоритмов на рынок труда России.

СПИСОК ЛИТЕРАТУРЫ

- [1]. Acemoglu D. et al., "AI and Jobs: Evidence from Online Vacancies," NBER Working Papers 28257, National Bureau of Economic Research, Inc., 2020
- [2]. Acemoglu D. et al., "Competing with Robots: Firm-Level Evidence from France," CEPR Discussion Papers 14971, C.E.P.R. Discussion Papers, 2020
- [3]. Acemoglu D., Restrepo P., "Automation and New Tasks: How Technology Displaces and Reinstates Labor." *Journal of Economic Perspectives* 33 (2), 2019, 3–30.
- [4]. Acemoglu D., Restrepo P., "Robots and Jobs: Evidence from US Labor Markets." *Journal of Political Economy*, 2020, 128(6): 2188–2244.
- [5]. Acemoglu D., Restrepo P., "The Wrong Kind of AI? Artificial Intelligence and the Future of Labour Demand." *Cambridge Journal of Regions, Economy and Society* 13(1), 2019, 25–35.
- [6]. Acemoglu D., Restrepo P., "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." *American Economic Review* 108(6), 2018, 1488-1542
- [7]. Aggarwal K. et al., "Has the future started? the current growth of artificial intelligence, machine learning, and deep learning," *Iraqi Journal for Computer Science and Mathematics*, vol. 3, no. 1, 2022, 115–123
- [8]. Autor, D. H. et al., "The Skill Content of Recent Technological Change: An Empirical Exploration." *The Quarterly Journal of Economics*, 118(4), 2003, 1279-1333,
- [9]. Brynjolfsson E. et al., "Machine Learning and Occupational Change." Unpublished manuscript, MIT, 2019
- [10]. Bughin J. et al., "Artificial Intelligence: The Next Digital Frontier?", McKinsey Global Institute, 2017

- [11]. Dinlersoz, E, Wolf Z., "Automation, Labor Share, and Productivity: Plant Level Evidence from U.S. Manufacturing," Working Papers 18-39, Center for Economic Studies, U.S. Census Bureau. 2018
- [12]. Felten E. et al, "The Effect of Artificial Intelligence on Human Labor: An Ability based Approach." *Academy of Management Proceedings*, Vol. 1, 2019
- [13]. Felten E. et al., "A Method to Link Advances in Artificial Intelligence to Occupational Abilities." *AEA Papers and Proceedings*, Vol. 108, 2018.
- [14]. Hosny A et al., "Artificial intelligence in radiology". *Nature Reviews. Cancer*. 18 (8), 2018, 500–510
- [15]. Quan, X.I.; Sanderson, J. Understanding the Artificial Intelligence Business Ecosystem. *IEEE Eng. Manag. Rev.*,46, 2018, 22–25
- [16]. Russell, S. "Human Compatible: Artificial Intelligence and the Problem of Control". United States: Viking. ISBN 978-0-525-55861-3, 2019.
- [17]. Webb, M., The impact of artificial intelligence on the labor market. Available at SSRN, 2020
- [18]. West D. M, "The Future of Work: Robots, AI, and Automation". Brookings Institution Press. 2018

ПРИЛОЖЕНИЯ

Приложение 1

Source	SS	df	MS	Number of obs	=	34,633
				F(153, 34479)	=	294.49
Model	46446.7687	153	303.573651	Prob > F	=	0.0000
Residual	35541.9871	34,479	1.03082999	R-squared	=	0.5665
				Adj R-squared	=	0.5646
Total	81988.7558	34,632	2.36742769	Root MSE	=	1.0153

ln_Vacanci~m	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
ln_Exp_AI_f	.2466427	.0132767	18.58	0.000	.2206199	.2726655

Приложение 2

Source	SS	df	MS	Number of obs	=	34,633
				F(153, 34479)	=	289.83
Model	38973.0726	153	254.725965	Prob > F	=	0.0000
Residual	30302.9497	34,479	.87888134	R-squared	=	0.5626
				Adj R-squared	=	0.5606
Total	69276.0223	34,632	2.00034714	Root MSE	=	.93749

ln_Non_AI_~s	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
ln_Exp_AI_f	.1832744	.0122592	14.95	0.000	.1592459	.2073028

Приложение 3

Source	SS	df	MS	Number of obs	=	34,633
				F(153, 34479)	=	106.04
Model	10392.2888	153	67.923456	Prob > F	=	0.0000
Residual	22084.6892	34,479	.640525805	R-squared	=	0.3200
				Adj R-squared	=	0.3170
Total	32476.978	34,632	.937773677	Root MSE	=	.80033

ln_AI_vaca~s	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
ln_Exp_AI_f	.2660015	.0104656	25.42	0.000	.2454885	.2865145

Приложение 4

Source	SS	df	MS	Number of obs	=	34,633
				F(154, 34478)	=	3.74
Model	5.7656401	154	.037439221	Prob > F	=	0.0000
Residual	344.981882	34,478	.010005855	R-squared	=	0.0164
				Adj R-squared	=	0.0120
Total	350.747522	34,632	.010127845	Root MSE	=	.10003

ln_share_Non_AI_vac	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
ln_Exp_AI_f	-.0245895	.0013081	-18.80	0.000	-.0271534	-.0220257

Приложение 5

Source	SS	df	MS	Number of obs	=	34,633
				F(154, 34478)	=	5.80
Model	12.0747525	154	.078407484	Prob > F	=	0.0000
Residual	466.089387	34,478	.013518458	R-squared	=	0.0253
				Adj R-squared	=	0.0209
Total	478.164139	34,632	.013807003	Root MSE	=	.11627

ln_share_AI_vac	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
ln_Exp_AI_f	.0341553	.0015204	22.46	0.000	.0311753	.0371354

Приложение 6

Source	SS	df	MS	Number of obs	=	8,544
				F(151, 8392)	=	70.86
Model	14792.9798	151	97.9667536	Prob > F	=	0.0000
Residual	11602.7392	8,392	1.38259523	R-squared	=	0.5604
				Adj R-squared	=	0.5525
Total	26395.719	8,543	3.08974821	Root MSE	=	1.1758

ln_Non_AI_vacancies	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
ln_Exp_AI_f	-.0308513	.0651025	-0.47	0.636	-.1584682	.0967656

Приложение 7

Source	SS	df	MS	Number of obs	=	8,544
				F(151, 8392)	=	12.28
Model	43.8745824	151	.290560149	Prob > F	=	0.0000
Residual	198.498191	8,392	.023653264	R-squared	=	0.1810
				Adj R-squared	=	0.1663
Total	242.372773	8,543	.02837092	Root MSE	=	.1538

ln_share_Non_AI_vac	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
ln_Exp_AI_f	-.0158839	.0085152	-1.87	0.062	-.0325758	.000808

Приложение 8

Source	SS	df	MS	Number of obs	=	8,544
				F(151, 8392)	=	36.24
Model	4423.46221	151	29.2944517	Prob > F	=	0.0000
Residual	6783.40229	8,392	.808317718	R-squared	=	0.3947
				Adj R-squared	=	0.3838
Total	11206.8645	8,543	1.31181839	Root MSE	=	.89906

ln_AI_vacancies	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
ln_Exp_AI_f	.5973393	.0497784	12.00	0.000	.4997613	.6949173

Приложение 9

Source	SS	df	MS	Number of obs	=	8,544
				F(151, 8392)	=	31.42
Model	3902.50129	151	25.8443794	Prob > F	=	0.0000
Residual	6903.13645	8,392	.822585372	R-squared	=	0.3612
				Adj R-squared	=	0.3497
Total	10805.6377	8,543	1.26485283	Root MSE	=	.90696

ln_share_AI_vac	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
ln_Exp_AI_f	.5860299	.0502158	11.67	0.000	.4875945	.684465

Приложение 10

Таблица 2. Результаты измерений только для фирм размещающих вакансии с задачами подходящими алгоритмам ИИ

	Growth of firm vacancies			
	AI vacancies	Non AI vacancies	Share of AI vacancies	Share of non AI vacancies
<i>Panel B: Additional data for firms set with nonzero AI vacancies (8544 observations)</i>				
Firm AI exposure	0.597*** (0.049)	-0.030 (0.065)	0.586*** (0.050)	-0.015*** (0.008)
<i>Fixed Effects:</i>				
Industry	✓	✓	✓	✓
Firm location	✓	✓	✓	✓
Firm size	✓	✓	✓	✓

Приложение 11

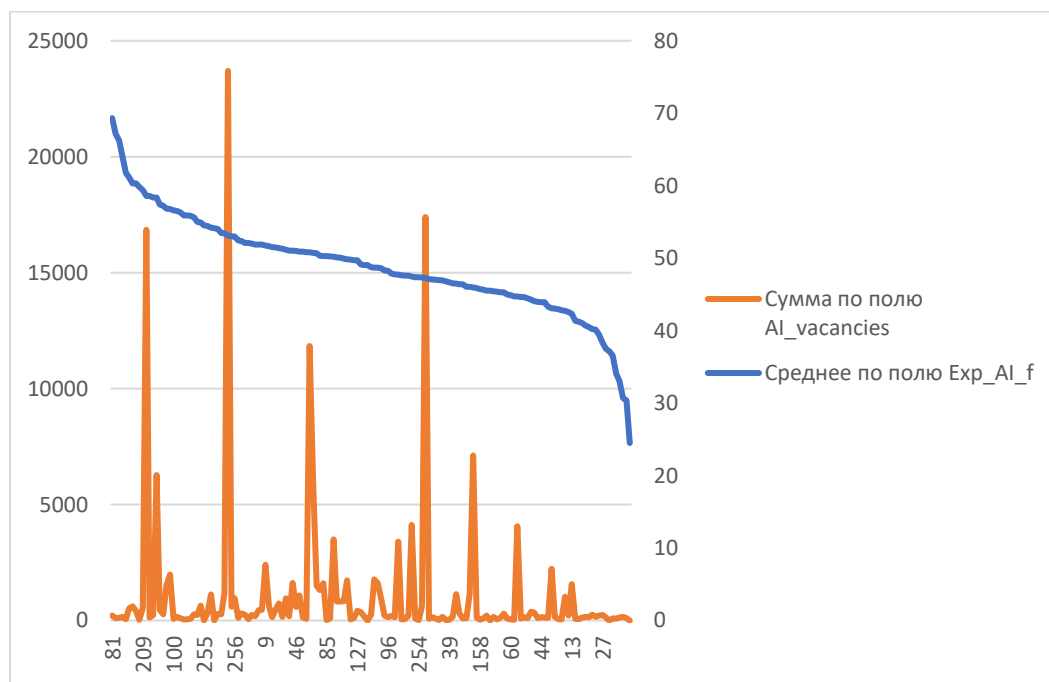


Рис. 1 Среднее значение $ExpAI_f$ по индустрии и количество вакансий ИИ в этой индустрии

Приложение 12

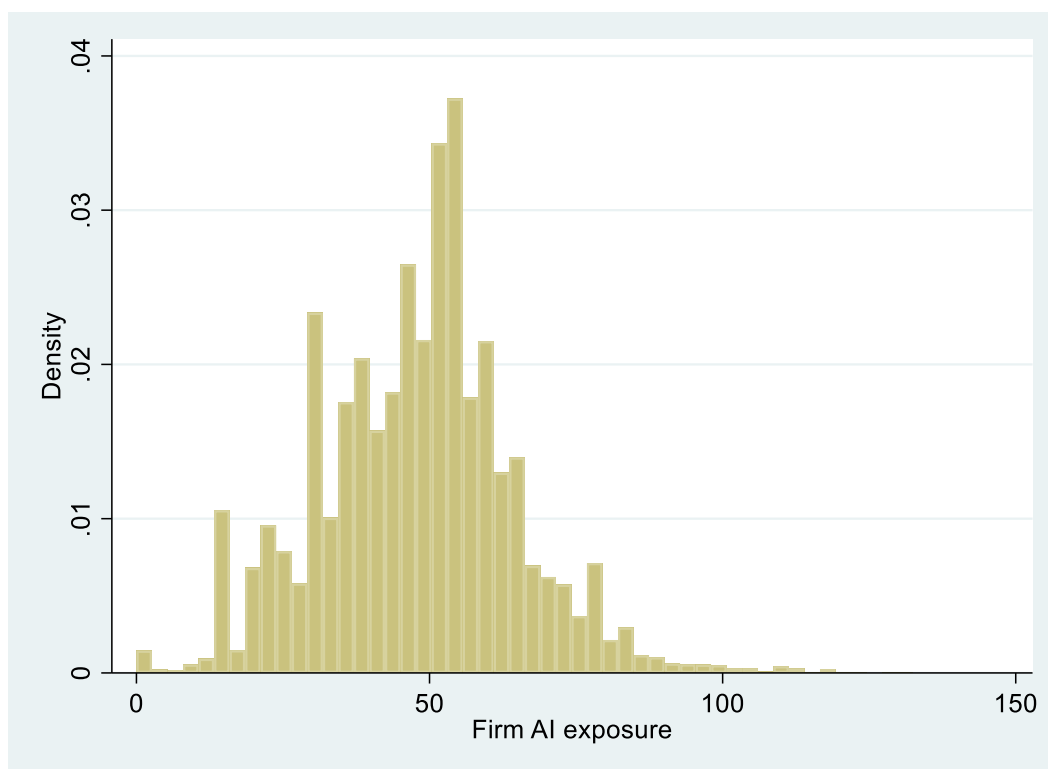


Рис. 2 Гистограмма распределения индекса влияния алгоритмов ИИ на фирму

Приложение 13

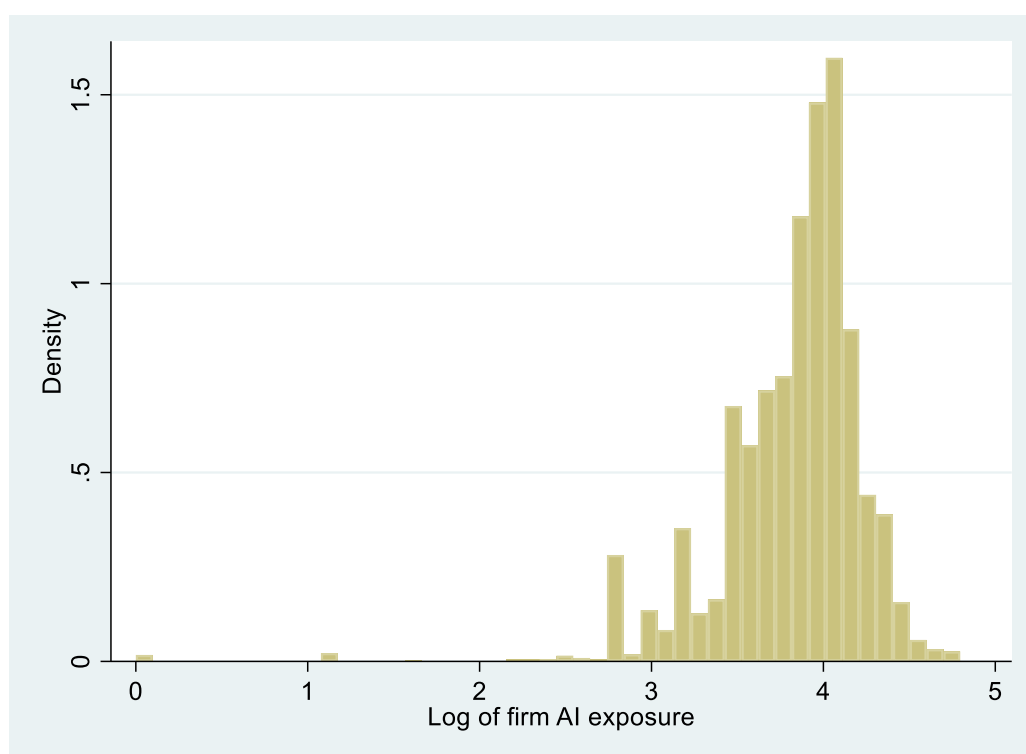


Рис. 3 Гистограмма распределения логарифма индекса влияния алгоритмов ИИ на фирму

Приложение 14

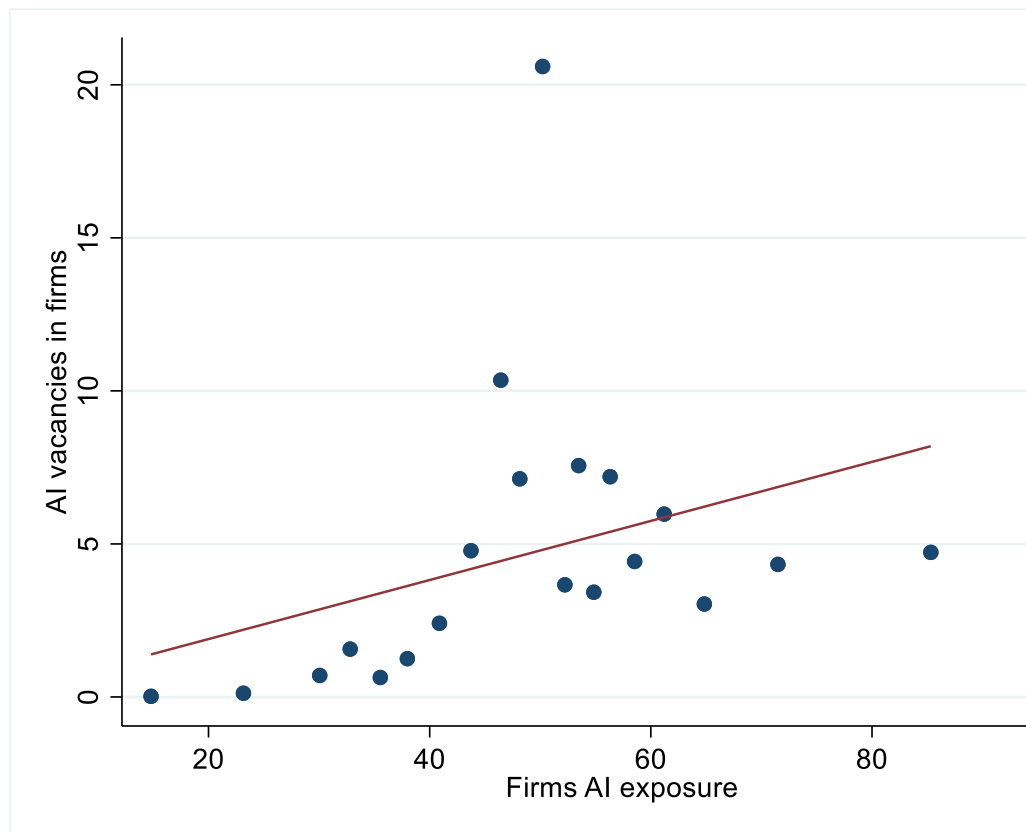


Рис. 4 Диаграмма рассеяния роста числа вакансий ИИ и $ExpAI_f$