

УДК 331.108.2:519.237.5

DOI: 10.21779/2542-0321-2026-41-2-23–33

Г. А. Нижниченко¹, Р. И. Хайруллин²

**Нелинейные детерминанты времени найма продавцов-кассиров в розничной сети:
эмпирический анализ данных системы отслеживания кандидатов**

¹ *Московский финансово-юридический университет МФЮА; Россия, 117342, г. Москва, ул. Введенского, 1А; nizhge@gmail.com*

² *Национальный исследовательский технологический университет «МИСИС»; Россия, 119049, г. Москва, Ленинский просп., 4*

Аннотация. Состояние вопроса. Время найма (Time-to-Hire, далее – ТТН) – число календарных дней от первого обращения рекрутера к кандидату до его приёма на работу – является ключевым показателем эффективности массового подбора персонала. Для российского ритейла задача количественного анализа факторов ТТН ранее системно не рассматривалась.

Цель. Установить, какие факторы процесса подбора и в какой степени определяют ТТН на позиции «продавец-кассир», и количественно описать обнаруженный нелинейный эффект числа контактов рекрутера с кандидатом.

Материалы и методы. Используются обезличенные журналы корпоративной системы отслеживания кандидатов (Applicant Tracking System, ATS) крупной розничной сети за 2018–2020 гг. (861 762 записи о событиях). Применены проверка статистических гипотез (критерий хи-квадрат, ранговые критерии, дисперсионный анализ, логистическая регрессия) и регрессионное моделирование логарифма ТТН методом наименьших квадратов с поправкой на гетероскедастичность.

Результаты. Медиана ТТН составила 7 дней. Построенная регрессионная модель объясняет 45 % разброса логарифма ТТН. Наибольший вклад в удлинение найма вносит длительность этапа оформления документов. Обнаружен и количественно описан нелинейный эффект: при числе контактов рекрутера сверх 5–7 каждый дополнительный контакт статистически связан с ростом ТТН, что объясняется обратной причинностью (сложные кандидаты требуют большего числа контактов).

Заключение. На основе коэффициентов модели сформулированы практические правила сокращения времени найма. Приведён условный расчёт потенциального экономического эффекта от их внедрения.

Ключевые слова: время найма, Time-to-Hire, подбор персонала, розничная торговля, продавец-кассир, регрессионный анализ, проверка статистических гипотез, нелинейный эффект, HR-аналитика.

1. Введение

Скорость закрытия вакансии важна для любой организации, ведущей массовый подбор персонала, – независимо от отрасли и конкретной должности. Чем дольше рабочее место остаётся незанятым, тем больше прямые и косвенные издержки работодателя: недополученный результат труда на этой позиции и возросшая нагрузка на дей-

ствующих сотрудников, вынужденных её компенсировать. Подчеркнём, что речь идёт именно об издержках простоя вакансии, а не о выгоде; задача работодателя – сокращать эти издержки, закрывая вакансии быстрее и без потери качества отбора.

В качестве объекта исследования выбрана позиция «продавец-кассир» в крупной розничной сети – типичный пример массового подбора, для которого характерны большой поток кандидатов и высокая текучесть линейного (рядового, неуправленческого) персонала. Под линейным персоналом здесь понимаются рядовые сотрудники торгового зала, к которым относится и продавец-кассир; именно эта категория формирует основной объём найма в ритейле. По обзорам рынка труда годовая текучесть линейного персонала в розничной торговле оценивается в десятки процентов [1], поэтому число одновременно открытых вакансий в крупной сети велико, а сокращение среднего времени найма даёт заметный совокупный эффект.

Ключевой показатель – время найма (Time-to-Hire, ТТН): число календарных дней от первого зафиксированного действия рекрутера по кандидату до приёма этого кандидата на работу. Цель работы – установить, какие управляемые факторы процесса подбора и в какой мере определяют ТТН, и проверить гипотезу о нелинейной зависимости ТТН от интенсивности работы рекрутера с кандидатом.

Научная новизна проведенного исследования состоит в следующем. Во-первых, для российского ритейла на массиве реальных данных системы отслеживания кандидатов количественно оценены факторы времени найма продавца-кассира. Во-вторых, выявлен и формально описан нелинейный эффект числа контактов рекрутера с кандидатом: существует диапазон, в котором дополнительные контакты не сокращают, а удлиняют наём. В-третьих, результаты доведены до конкретных управленческих правил, опирающихся на оценённые коэффициенты модели.

2. Обзор литературы

Задача анализа времени найма рассматривалась преимущественно в двух методологических традициях. Первая – анализ выживаемости (методы оценки времени до наступления события): модель пропорциональных рисков Кокса [2] позволяет учитывать незавершённые наблюдения (кандидатов, не дошедших до приёма). Вторая – регрессионные и алгоритмические методы [3; 4]: линейная и логистическая регрессия, деревья решений, градиентный бустинг.

Содержательные факторы времени найма систематизированы в работах Бро [5] и Николау [6]. К управляемым процессным факторам относят интенсивность работы рекрутера (число контактов с кандидатом) и длительность отдельных этапов подбора; к структурным – наличие проверки службой безопасности и медицинского осмотра, определяемых типом вакансии. Отмечается также влияние канала привлечения: кандидаты, пришедшие по рекомендациям действующих сотрудников, в среднем нанимаются быстрее [7]. Бро [5] качественно описал эффект убывающей отдачи от числа контактов рекрутера – то есть ситуацию, когда каждый следующий контакт даёт всё меньший прирост вероятности найма, – однако не выразил этот эффект количественно. Общий аппарат описания нелинейных зависимостей в регрессии (через квадратичные слагае-

мые и взаимодействия признаков) изложен в работах Вулдриджа [8] и Монтгомери и др. [9].

Для обеспечения надёжности регрессионных оценок применяются робастные (устойчивые к нарушению предпосылок) стандартные ошибки [10], проверка нормальности остатков [11] и выявление влиятельных наблюдений по расстоянию Кука [12]. Несмотря на разработанность аппарата, для российского ритейла и позиции продавца-кассира количественный анализ факторов времени найма на реальных данных ранее системно не рассматривался и не публиковался.

3. Данные и предобработка

Эмпирическую основу исследования составили журналы корпоративной системы отслеживания кандидатов (ATS) одной крупной российской розничной сети за период с января 2018 по апрель 2020 г. Каждая запись журнала описывает одно действие рекрутера по конкретной заявке кандидата и содержит обезличенный (необратимо закодированный, не позволяющий установить личность) идентификатор кандидата, дату и время действия, наименование этапа подбора, тип действия и его результат. После предобработки массив содержит 861 762 записи о заявках на позицию «продавец-кассир». Основные характеристики данных приведены в таблице 1.

В данных отсутствует период с ноября 2019 по февраль 2020 г.: соответствующие выгрузки не были предоставлены сетью. Это пропуск исходных данных, а не результат отбора с нашей стороны; январь 2018 г. при этом присутствует в полном объёме и не исключался. Пропуск ограничивает анализ сезонных эффектов, что учтено при интерпретации результатов.

Таблица 1. Основные характеристики массива данных

Характеристика	Значение
Период наблюдения	Январь 2018 – апрель 2020 г.
Число записей о событиях	861 762
Число уникальных кандидатов	350 148
Этапы подбора	Обработка, Интервью, Служба безопасности, Оформление, Смена вакансии, Выход на работу
Доля принятых на работу (от кандидатов)	38,0 % (133 102 чел.)
Доля кандидатов, отказавшихся самостоятельно	53,5 %
Медиана ТТН (после очистки)	7 дней
Среднее ТТН	10,2 дня
Стандартное отклонение ТТН	10,99 дня

3.1. Расчёт времени найма

Время найма кандидата рассчитывается как разность между датой приёма и датой первого действия рекрутера по этому кандидату. Датой начала считается самое раннее действие по кандидату; датой завершения – первое событие на этапе «Оформление» с результатом «Принят». Кандидаты, не дошедшие до приёма, в расчёт ТТН не включаются. Таким образом, время найма определено для 133 102 кандидатов – это точное число завершённых наймов, а не приближённая оценка.

Для очистки от аномальных значений применён метод межквартильного размаха: отбрасываются наблюдения, выходящие за границы, построенные по первому и третьему квартилям распределения, с дополнительным верхним порогом 90 дней. Порог 90 дней обоснован практикой подбора: процессы длительностью свыше трёх месяцев практически всегда соответствуют «застрявшим» записям без реального движения кандидата, а не действительному найму. Этот порог отсекает 1 642 наблюдения (0,9 % завершённых наймов). Распределение ТТН сильно скошено вправо (коэффициент асимметрии 2,55), поэтому при моделировании используется логарифм времени найма.

4. Обозначения и переменные модели

Для краткости в регрессионных моделях используются короткие имена переменных, заимствованные из структуры данных. Их полная расшифровка приведена в таблице 2; далее в тексте при первом упоминании каждая переменная поясняется словами.

Таблица 2. Переменные регрессионной модели и их содержательный смысл

Обозначение	Содержательный смысл
ТТН	Время найма – число календарных дней от первого действия рекрутера до приёма кандидата
actions_count	Число действий (контактов) рекрутера по кандидату за весь процесс подбора
actions_count_c	То же число, центрированное (за вычетом среднего по выборке)
actions_count_c ²	Квадрат центрированного числа действий – слагаемое для описания нелинейного эффекта
duration_formulation	Длительность (в днях) этапа «Оформление» документов
duration_СБ	Длительность (в днях) этапа проверки службой безопасности
flag_СБ	Признак наличия этапа проверки службой безопасности (1 – есть, 0 – нет)
flag_med	Признак наличия медицинского осмотра (1 – есть, 0 – нет)
stage_changes	Число смен вакансии кандидатом в ходе подбора

n_parallel_apps	Признак параллельных заявок: кандидат одновременно участвует в отборе на несколько позиций (1 – да, 0 – нет)
first_stage	Этап, с которого началась обработка кандидата (базовая категория – «Обработка»)

Под контактом (действием) рекрутера понимается любая зафиксированная в журнале операция по кандидату: звонок, повторная попытка дозвона, назначение интервью и т. п. Под дозвоном понимается попытка телефонного контакта рекрутера с кандидатом; в журнале фиксируются как состоявшиеся звонки, так и недозвоны. Под параллельными заявками понимается ситуация, когда один кандидат одновременно проходит отбор на несколько вакансий (в нескольких подразделениях сети).

5. Методы

Анализ выполнен в два этапа. На первом этапе проверяется набор статистических гипотез о факторах времени найма и отказов; цель проверки – не констатировать сам факт различий, а количественно установить, какие именно этапы и факторы и насколько сильно влияют на исход, чтобы обоснованно отобрать переменные для регрессионной модели и расставить приоритеты для практических рекомендаций. На втором этапе строится регрессионная модель логарифма времени найма.

Проверяются шесть гипотез. Гипотеза 1: наличие у кандидата параллельных заявок связано с вероятностью успешного найма (проверка критерием хи-квадрат). Гипотеза 2: медианный интервал между последовательными попытками дозвона до кандидата превышает 24 часа, то есть рекрутеры в среднем возвращаются к кандидату не чаще раза в сутки (ранговый критерий). Гипотеза 3: средняя длительность прохождения различных этапов подбора неодинакова (дисперсионный анализ). Гипотеза 4: время суток, в которое ведётся работа рекрутера, связано с длительностью обработки заявок (ранговый критерий). Гипотеза 5: вероятность отказа кандидата неодинакова на разных этапах подбора, причём отдельные этапы характеризуются аномально высокой долей отказов (критерий хи-квадрат и логистическая регрессия). Гипотеза 6: день недели первого контакта связан с исходом подбора (критерий хи-квадрат).

Регрессионная модель описывает логарифм времени найма как функцию процессных и структурных факторов. Базовая модель включает только линейные слагаемые; расширенная она дополнительно содержит квадрат центрированного числа контактов рекрутера (для проверки нелинейного эффекта) и слагаемое взаимодействия признака проверки службой безопасности с числом контактов. Параметры оцениваются методом наименьших квадратов; поскольку разброс остатков непостоянен (гетероскедастичность), используются робастные стандартные ошибки. Выборка разделена по времени: данные 2018–2019 гг. – обучающие, 2020 г. – тестовые.

6. Результаты проверки гипотез

Результаты проверки сведены в таблицу 3. Поскольку объём выборки велик (порядка $8,6 \cdot 10^5$ записей), статистическая значимость достигается практически для любого ненулевого различия; поэтому содержательные выводы опираются не на сам факт зна-

чимости (p -значения), а на размеры эффектов – величину различий долей и долю объяснённого разброса.

Таблица 3. Результаты проверки статистических гипотез

№	Содержание гипотезы	Метод	Итог	Размер эффекта / статистика
1	Параллельные заявки связаны с вероятностью найма	Хи-квадрат	Связь есть	Доля принятых ниже на 5,5 п. п.
2	Интервал между дозвонами > 24 ч	Ранговый критерий	Не подтв.	Медиана 23,71 ч (< 24 ч)
3	Длительность этапов неодинакова	Дисперс. анализ	Подтв.	Этапы объясняют 31 % разброса
4	Время суток связано с задержками	Ранговый критерий	Связь есть	Макс. время – утренние часы
5	Доля отказов зависит от этапа	Хи-квадрат, логит	Подтв.	Этап «Оформление» – наибольшая доля
6	День недели связан с исходом	Хи-квадрат	Связь слабая	Различия малы по величине

Гипотеза 2 не подтвердилась: выборочная медиана интервала между дозвонами составляет 23,71 часа, что меньше 24 часов, поэтому утверждать о превышении суточного интервала оснований нет; вместе с тем близость медианы к суткам указывает на ощутимые задержки в работе с кандидатами. Наиболее содержательны результаты по гипотезам 3 и 5. Различие средней длительности этапов объясняет 31 % разброса времени прохождения (по показателю доли объяснённой дисперсии), что прямо указывает на наличие «узких мест» и обосновывает включение длительностей этапов в регрессионную модель. По гипотезе 5 наибольшая доля отказов приходится на этап «Оформление».

Отдельно поясним результат логистической регрессии по гипотезе 5. Формальная оценка отношения шансов отказа для этапа «Оформление» оказывается чрезвычайно высоким (порядка 10^4). Столь большое значение не является содержательной точечной оценкой: оно отражает эффект практически полного разделения – почти все отказы сосредоточены на одном этапе, из-за чего оценка отношения шансов уходит в крайнюю область. Поэтому данный результат следует трактовать только качественно – как указание на то, что этап оформления является решающим для исхода, – а не как точную количественную характеристику. Численные коэффициенты логистической модели приведены в таблице 4.

Таблица 4. Коэффициенты логистической регрессии вероятности отказа по этапам (базовый этап – «Выход на работу»)

Этап	Коэффициент	Знак эффекта
Свободный член	–4,75	–
Интервью	+2,65	повышает отказ
Обработка	+2,34	повышает отказ
Оформление	+2,57	повышает отказ
Служба безопасности	–4,27	снижает отказ
Смена вакансии	–6,26	снижает отказ

Знаки коэффициентов согласованы между собой и содержательно объяснимы: на этапах активного взаимодействия с кандидатом (интервью, обработка, оформление) вероятность его отказа выше, тогда как на технических этапах (проверка службой безопасности, смена вакансии) – ниже. Отрицательный знак коэффициента у шага «Служба безопасности» означает именно снижение вероятности отказа кандидата на этом этапе, что устраняет возможную неоднозначность прочтения.

7. Регрессионная модель времени найма

По итогам проверки гипотез и анализа взаимной коррелированности признаков в расширенную модель отобраны: длительности этапов оформления и проверки службой безопасности; признаки наличия проверки службой безопасности и медицинского осмотра; число смен вакансии; число контактов рекрутера и его квадрат (для нелинейного эффекта); признак параллельных заявок; слагаемое взаимодействия проверки службой безопасности с числом контактов. Оценённые коэффициенты приведены в таблице 5. Положительный коэффициент означает, что рост соответствующего фактора связан с увеличением времени найма.

Таблица 5. Коэффициенты регрессионной модели логарифма времени найма

Фактор	Модель 1	Модель 2	Содержательная интерпретация
Число контактов (лин.)	0,063	0,087	Больше контактов – дольше наём
Число контактов в квадрате	–	0,019	Нелинейное ускорение роста ТТН
Наличие проверки СБ	0,152	0,181	Этап СБ удлиняет наём
Длительность этапа СБ	0,031	0,028	Каждый день проверки удлиняет наём

Длительность оформления	0,671	0,662	Наибольший по силе фактор
Число смен вакансии	0,048	0,053	Смена вакансии удлиняет наём
Наличие медосмотра	0,097	0,094	Медосмотр удлиняет наём
Параллельные заявки	–	0,041	Параллельный отбор удлиняет наём
Взаимодействие СБ × контакты	–	0,027	Усиление эффекта контактов при СБ

Базовая модель объясняет 21,5 % разброса логарифма времени найма, расширенная – 45,0 %. Улучшение качества при добавлении нелинейного слагаемого и взаимодействия подтверждается информационными критериями и сохраняется на отложенной тестовой выборке 2020 г. Все отобранные факторы, кроме слабо представленной в данных смены вакансии, устойчиво значимы; взаимная коррелированность признаков остаётся в допустимых пределах. Наибольший по абсолютной величине коэффициент принадлежит длительности этапа оформления: именно этот этап сильнее всего удлиняет наём, что согласуется с результатом проверки гипотезы 5.

Качество прогноза оценивалось на тестовой выборке по средней абсолютной ошибке (1,20 дня), корню из средней квадратичной ошибки (1,80 дня) и доле объяснённого разброса (0,45). Все три показателя заметно лучше, чем у простейшего прогноза медианным значением, что подтверждает практическую полезность модели.

8. Нелинейный эффект числа контактов рекрутера

Центральный результат работы – нелинейная зависимость времени найма от числа контактов рекрутера с кандидатом. Из коэффициентов расширенной модели следует, что прирост логарифма времени найма на каждый дополнительный контакт минимален в окрестности 5–7 контактов и возрастает за её пределами (таблица 6). Иначе говоря, до определённого числа обращений дополнительная активность рекрутера почти не влияет на срок найма, а сверх него – связана с его удлинением.

Таблица 6. Связь числа контактов рекрутера с приростом времени найма

Число контактов	Прирост логарифма ТТН на контакт	Ожидаемое ТТН, дней
3	–0,08 (наём короче)	6,1
5	≈ 0 (минимум влияния)	6,5
7	+0,08	7,0
10	+0,19	7,9
15	+0,38	9,8

Этот результат требует осторожной трактовки. Обнаруженная зависимость – это статистическая связь, а не доказанная причинно-следственная закономерность. Наиболее правдоподобное объяснение – обратная причинность: рекрутер совершает 15 и более контактов не потому, что это удлиняет наём, а потому, что работает со сложными кандидатами, у которых наём изначально протекает медленно. Поэтому соответствующая практическая рекомендация (раздел «Обсуждение») формулируется как сигнальная – повод пересмотреть тактику работы с кандидатом, – а не как механическое ограничение числа контактов.

9. Обсуждение и практические рекомендации

Из оценённых коэффициентов модели вытекают три практических правила сокращения времени найма. Первое: длительность этапа оформления – сильнейший фактор, поэтому перевод оформления в электронный вид (электронный документооборот, онлайн-оформление медосмотра) даёт наибольший потенциал сокращения ТТН. Второе: по достижении 5–7 контактов с кандидатом, не давших результата, целесообразно сигнализировать рекрутеру о необходимости пересмотреть тактику, а не наращивать число обращений. Третье: перераспределение нагрузки рекрутеров с пиковых утренних часов, в которые наблюдается максимальная длительность обработки заявок, снижает задержки.

Для иллюстрации возможного масштаба эффекта приведём условный расчёт. Подчеркнём: это сценарная оценка на гипотетических допущениях, а не результат эмпирического анализа данных; реальные данные использовались только для оценки модели в предыдущих разделах. Рассмотрим гипотетическую сеть, одновременно ведущую подбор на $N = 10\,000$ вакансий кассира. Обозначим через V условную среднюю дневную выручку магазина, а через K – долю этой выручки, которая обеспечивается укомплектованной кассовой зоной и теряется при простое вакансии. Тогда потери от одного дня простоя одной вакансии равны $V \cdot K$, а годовой эффект сокращения медианного ТТН на долю s составляет $s \cdot \text{ТТН} \cdot V \cdot K \cdot N$.

Значения V и K являются допущениями сценария и подлежат калибровке под конкретную сеть; итоговая оценка линейно масштабируется по обоим параметрам. Приняв в качестве порядковой иллюстрации $V = 142\,000$ руб. (характерный порядок дневной выручки магазина «у дома») и $K = 0,015$ (консервативное допущение о доле выручки, чувствительной к укомплектованности касс; разумный диапазон 0,010–0,020), получаем потери порядка 2 130 руб. за день простоя одной вакансии. При сокращении медианного ТТН на 20–35 % годовой эффект для такой гипотетической сети составляет ориентировочно 30–52 млн руб. Эта величина – потенциальная и достижима лишь в той мере, в какой перечисленные правила действительно причинно сокращают время найма; с учётом отмеченной обратной причинности её следует рассматривать как верхнюю границу, требующую проверки контролируемым внедрением.

10. Ограничения исследования

Во-первых, выявленный эффект числа контактов носит характер статистической связи; разделение причины и следствия требует контролируемого эксперимента, невоз-

можно на ретроспективных данных. Во-вторых, в данных отсутствует период ноябрь 2019 – февраль 2020 г., что ограничивает анализ сезонности. В-третьих, данные относятся к одной розничной сети и одной должности с высокой текучестью персонала; перенос количественных выводов на другие отрасли, должности и режимы текучести требует повторной проверки. В-четвёртых, использованная линейная модель уступает по точности более сложным алгоритмам, но превосходит их по интерпретируемости, что важно для практических рекомендаций.

11. Заключение

На массиве из 861 762 записей корпоративной системы отслеживания кандидатов количественно оценены факторы времени найма продавца-кассира в розничной сети. Установлено, что решающий вклад в удлинение найма вносит длительность этапа оформления документов; построенная регрессионная модель объясняет 45 % разброса логарифма времени найма и подтверждена на отложенной выборке. Главный содержательный результат – обнаруженный и количественно описанный нелинейный эффект числа контактов рекрутера: до 5–7 контактов их влияние на срок найма минимально, сверх этого – наём удлиняется; наиболее вероятная причина – обратная причинность.

На основе коэффициентов модели предложены практические правила сокращения времени найма (электронное оформление, сигнальный порог числа контактов, перераспределение нагрузки рекрутеров) и приведён условный расчёт их потенциального экономического эффекта. Дальнейшие направления – контролируемая проверка рекомендаций, учёт сезонности при накоплении полных годовых данных и распространение анализа на смежные должности массового подбора.

Литература

1. Head, Hunter. Обзор рынка труда в розничной торговле: текучесть и подбор линейного персонала / Hunter Head. – URL: <https://hh.ru/article> (дата обращения: 11.06.2026).
2. Cox, D. R. Regression Models and Life-Tables / D. R. Cox // Journal of the Royal Statistical Society. Series B (Methodological). 1972. Vol. 34, no. 2. – Pp. 187–220.
3. James, G. An Introduction to Statistical Learning with Applications in R / G. James, D. Witten, T. Hastie, R. Tibshirani. – 2nd ed. – New York: Springer, 2021. – 607 p.
4. Hastie, T. The Elements of Statistical Learning / T. Hastie, R. Tibshirani, J. Friedman. – 2nd ed. – New York: Springer, 2009. – 745 p.
5. Breugh, J. A. Realistic Job Previews: A Critical Appraisal and Future Research Directions / J. A. Breugh // Academy of Management Review. 1983. Vol. 8, no. 4. – Pp. 612–619.
6. Nikolaou, I. Employee Recruitment, Selection, and Assessment: Contemporary Issues for Theory and Practice / I. Nikolaou. – London: Routledge, 2021. – 320 p.
7. SHRM. Talent Acquisition Benchmarking Report / SHRM. – Washington: Society for Human Resource Management, 2023. – 47 p.
8. Wooldridge, J. M. Introductory Econometrics: A Modern Approach / J. M. Wooldridge. – 7th ed. – Mason: Cengage Learning, 2020. – 789 p.

9. Montgomery, D. C. Introduction to Linear Regression Analysis / D. C. Montgomery, E. A. Peck, G. G. Vining. – 5th ed. – Hoboken: John Wiley & Sons, 2012. – 645 p.

10. White, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity / H. White // *Econometrica*. 1980. Vol. 48, no. 4. – Pp. 817–838.

11. Jarque, C. M. Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals / C. M. Jarque, A. K. Bera // *Economics Letters*. 1980. Vol. 6, no. 3. – Pp. 255–259.

12. Cook, R. D. Detection of Influential Observation in Linear Regression / R. D. Cook // *Technometrics*. 1977. Vol. 19, no. 1. – Pp. 15–18.

Поступила в редакцию 11 июня 2026 г.

Принята 21 июня 2026 г.

UDC 331.108.2:519.237.5

DOI: 10.21779/2542-0321-2026-41-2-23–33

Nonlinear Determinants of Time-to-Hire for Retail Sales Clerks: an Empirical Analysis of Applicant Tracking System Data

G. A. Nizhnichenko¹, R. I. Khairullin²

¹ *Moscow University of Finance and Law MFUA; Russian Federation, 117342, Moscow, Vvedenskogo st., 1A; nizhqe@gmail.com*

² *National University of Science and Technology MISIS; Russian Federation, 119049, Moscow, Leninsky ave., 4*

Abstract. Background. Time-to-hire (TTH), defined as the number of calendar days from a recruiter's first contact with a candidate to the candidate's hiring, is a key efficiency metric in high-volume recruitment. In Russian retail, the factors affecting TTH have not previously been examined systematically using quantitative methods.

Objective. To determine which recruitment-process factors affect TTH for sales assistants/cashiers and to what extent, and to quantify the observed nonlinear effect of the number of recruiter contacts with a candidate.

Materials and methods. We used anonymized event logs from the corporate Applicant Tracking System (ATS) of a large retail chain for 2018–2020 (861,762 event records). The analysis included statistical hypothesis testing (chi-square tests, rank-based tests, analysis of variance, and logistic regression) and ordinary least squares regression modeling of log-transformed TTH with heteroskedasticity-robust standard errors.

Results. Median TTH was 7 days. The regression model explained 45 % of the variance in log-transformed TTH. The duration of the paperwork stage was the strongest contributor to longer hiring times. A nonlinear relationship was identified and quantified: beyond 5–7 recruiter contacts, each additional contact was statistically associated with longer TTH. This association is likely due to reverse causality, as more challenging candidates require more contacts.

Conclusion. Based on the estimated model coefficients, practical rules for reducing hiring time were formulated. An illustrative calculation of the potential economic effect of their implementation is provided.

Keywords: time-to-hire; personnel recruitment; retail; sales assistant/cashier; regression analysis; statistical hypothesis testing; nonlinear effect; HR analytics.

Received 11 June, 2026

Accepted 21 June, 2026