

Трек №2. Уволить нельзя нанять.

Оценка влияния пандемии на рынок труда.

Барамия Никита, Кагинян Маргарита, Комарова Анна
Экономический факультет МГУ им. М.В. Ломоносова

20 декабря 2020 г.

Содержание

1 Введение	1
2 Анализ данных	2
3 Эмпирический подход	6
4 Результаты	7

1 Введение

В данной работе мы приводим результаты нашего исследования поведения на рынке труда зарегистрированных безработных в период пандемии коронавируса и обращаем внимание органов власти на возможные места оптимизации порядка назначения выплат и присвоения статуса "безработный".

Мировая практика подтверждает эмпирическую закономерность: чем больше размер пособия по безработице и чем дольше продолжительность выплат, тем дольше люди остаются в статусе безработного, поскольку это увеличивает их нетрудовой доход и резервную заработную плату. Поэтому в данном исследовании в качестве одной из гипотез мы выдвигаем следующую: дополнительные выплаты, введенные постановлением Правительства Российской Федерации, на каждого несовершеннолетнего ребенка в размере 3000 рублей стимулируют рост безработицы, вызванный желанием получить дополнительный доход.

Данная гипотеза подтвердилась в ходе эконометрического исследования, описанного ниже, и мы можем сказать, что увеличение суммы дополнительных выплат снижает вероятность выхода на рынок труда в конкретный момент времени на 2%. Факт подтверждения этой гипотезы говорит о том, что введение подобных выплат

искажает стимулы безработных к трудоустройству и, следовательно, бюджет распределяется не в целевые группы населения.

Также мы предполагаем, что одним из решающих факторов, снижающих вероятность трудоустройства (или же увеличивая продолжительность пребывания в статусе безработного), является характеристика отсутствия трудовой деятельности у заявителя. Это большая социальная проблема, поскольку самая большая доля людей без опыта (57%) – это люди в возрасте от 18 до 30 лет, причем у 25% есть дети (и разумно предположить, что это в основном дети до 10 лет). Причем эту группу можно охарактеризовать, как высоко мобильную, поскольку 85.3% из этой группы готовы к переквалификации (против 80% в группе 31-49 лет и 77.8% в группе 41-50 лет). Мы считаем, что эта группа особенно нуждается в государственной поддержке со стороны служб занятости.

2 Анализ данных

Для нашего исследования мы выбрали набор сведений о отдельных гражданах, обратившихся в электронной форме за пособием по безработице в период с 9 апреля по 31 октября 2020 года¹ по нескольким причинам: во-первых, выдвинутые нами гипотезы требовали данных и выплатах, полученных безработными, которые отсутствуют в наборе данных 1; во-вторых, фильтр заявлений, использованный нами, позволяет делать такие же выводы, как и по официальному реестру (безработные-1). Предварительно была произведена очистка по status: далее рассматривались только заявления со статусами APPROVED и EMPLOYED, таким образом, в нашей выборке присутствуют только официально зарегистрированные безработные и те из них, кто трудоустроился на момент выгрузки данных. Помимо этого были удалены присутствующие в выборке дубликаты (17 тыс.): итого осталось чуть больше 2.5 млн. наблюдений. При очистке данных от пропущенных значений осталось 513 тысяч наблюдений, так что будет разумно сравнить распределения ряда характеристик, чтобы более обоснованно говорить о репрезентативности итоговой выборки и будущей генерализации результатов. На рисунках 1, 2, 3 мы можем наблюдать сохранения баланса по ряду важных демографических характеристик.

Рисунок 1 демонстрирует распределение по возрастам. Предполагать линейную зависимость от возраста кажется нелогичным, так что для нашей будущей модели мы создадим бинарные переменные на основе имеющихся данных: принадлежность к группам 18-30, 31-40, 41-50. Были обнаружены 78 человек в возрасте 15-17 лет. Мы решили удалить их из выборки, так как они вне сферы наших исследовательских вопросов, да и могли бы потребовать более кропотливой очистки данных: делать выводы на основе 78 наблюдений из 400 тысяч не представляется возможным. Аналогично поступим с людьми старше 60 лет. Таким образом, мы будем сравнивать ве-

¹условное название - безработные-2. Источник: Роструд; обработка: Роструд, Инфраструктура научно-исследовательских данных, АНО "ЦПУР 2020 (набор данных получен в рамках хакатона PandemicDataHack 18-20 декабря 2020)

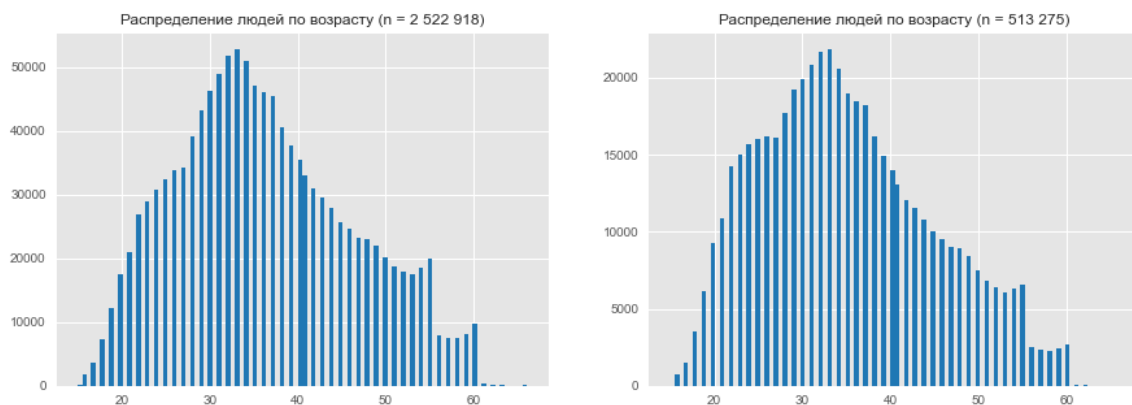


Рис. 1. Сохранение распределения людей по возрасту

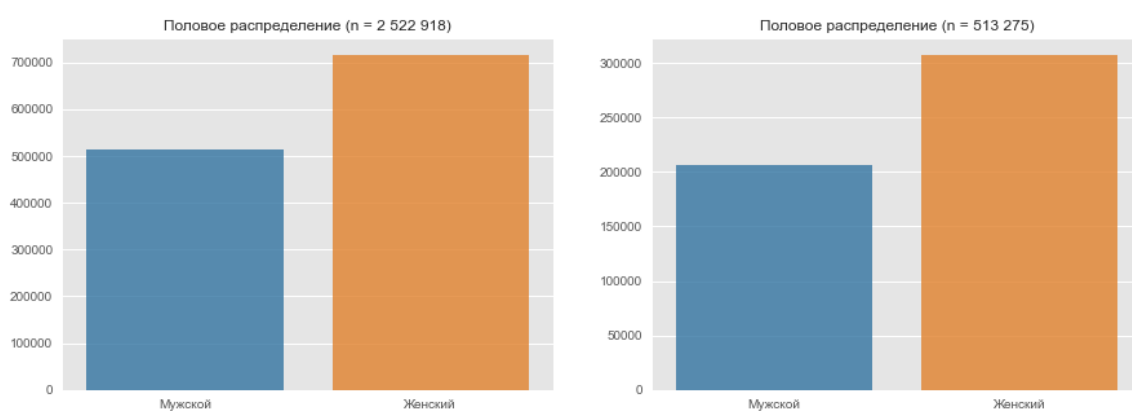


Рис. 2. Сохранение распределения людей по полу

роятность людей из выше перечисленных групп выйти из безработицы относительно группы 51-60 лет. На рисунке 2 мы видим хороший баланс между полами, что позволит обеспечить качественный контроль на половую принадлежность и узнать, кому было проще во время пандемии найти работу. На рисунке 3 мы видим малое количество людей с незаконченным высшим образованием, так что в качестве контроля будут участвовать только наличие среднего профессионального образования и высшее образование и будет оцениваться их эффект по сравнению с двумя оставшимися группами.

Помимо описанных выше мы также принимаем во внимание следующие характеристики:

- `did_not_work` (признак отсутствия трудовой деятельности) – здравый смысл подсказывает, что людям, ни разу не работавшим, сложнее трудоустроиться. В нашей итоговой выборке таких людей 24%.
- `social_status_has_child` (наличие у заявителя детей несовершеннолетнего возраста) – здесь возможны два противодействующих эффекта: с одной стороны, ребёнок требует внимания, поэтому издержки на поиск работы повышаются,

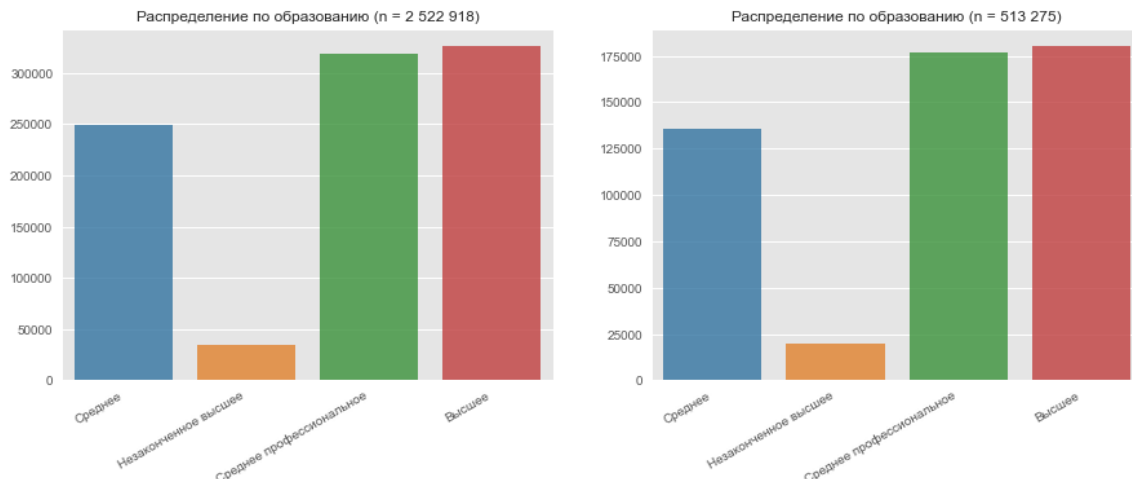


Рис. 3. Сохранение распределения людей по образованию

с другой стороны, наличие детей повышает расходы, так что возможен рост стимулов к трудоустройству. В нашей выборке 38% таких людей.

- `cv_experience` (опыт работы) – больше опыта работы, выше шансы на трудоустройство. Средний опыт соответствует 5 годам.
- `cv_salary` (ожидаемая заработная плата) – выше ожидания, ниже шансы трудоустройства, поскольку список вакансий с ростом ожиданий будет сокращаться. Средние ожидания соответствуют 33 тысячам рублей.
- `cv_retrainingcapability` (готовность к переобучению) – свидетельствует о большей целеустремлённости гражданина. 78% указали, что готовы к переобучению.
- `sum_social_benefits`, `sum_child_social_benefits` – чем выше пособия, тем меньше стимулы для активного поиска работы
- `work_2019_7_2020_9` – данная переменная была создана нами на основе переменных `work_in_2020_7 ... work_in_2020_9`, означая наличие работы в один из данных месяцев. 75% нашей выборки имели работу в один из месяцев второго полугодия 2019 года или в первых трёх кварталах 2020.
- `last_salary` – также созданная нами переменная, аналогично способу с переменной `work_2019_7_2020_9`, говорящая о последнем квартальном уровне зп работника.

Также мы использовали ряд данных из внешних источников.

В качестве контроля на особенности региональных рынков труда и выделения той части официальной безработицы, которая может быть вызвана именно пандемией коронавируса, мы контролируем на средний уровень безработицы в каждом

регионе за прошедшие 3 года 2017-2019. Данные опубликованы федеральной службой официальной статистики ЕМИСС ².

Дополнительно мы выдвигаем гипотезу о том, что в регионах, где большая доля экономически активного населения занята в отраслях, наиболее пострадавших от пандемии ³, во-первых, будет в среднем выше уровень безработицы, во-вторых, сложнее будет процесс трудоустройства, в связи с необходимостью времени на восстановление производства. Поэтому мы составили соответствующую переменную на основе данных ЕМИСС по среднегодовой численности занятых по видам экономической деятельности ⁴. Итого к наиболее пострадавшим отраслям мы отнесли следующие: деятельность в области культуры и спорта, организация досуга и развлечений, деятельность гостиниц и предприятий общественного питания,

При первичном анализе данных мы выявили следующие закономерности:

Пол	Наличие детей	Доля трудоустроенных
Мужчины	Нет	16%
	Да	13.5%
Женщины	Нет	14.9%
	Да	10.6%

Таблица 1. Соотношение семейного статуса и статуса безработного

Проанализировав данные из таблицы 1, можно предположить, что наличие детей в семье отрицательно влияет на вероятность выхода на рынок труда, поскольку дети требуют ухода и большое количество свободного времени. Однако мы выдвигаем гипотезу о том, что именно получение дополнительных выплат, введенных для безработных с несовершеннолетними детьми, может являться сдерживающим фактором от выхода на рынок труда и смещать истинное влияние наличия детей на вероятность трудоустройства.

Подводя итог, мы выдвигаем следующие гипотезы:

1. Наличие детей стимулирует человека к более активному поиску работы, что выражается в более высокой вероятности выхода на работу;
2. Вероятность выхода на работу для женщин ниже, чем для мужчин;
3. Люди с высшим образованием более вероятно найдут работу, чем люди со средним или неоконченным высшим;
4. Люди из регионов с более высокой безработицей будут иметь меньшую вероятность выхода по сравнению с регионами более низких уровней.

²<https://www.fedstat.ru/indicator/33446>

³Перечень министерства экономического развития: <https://мойбизнес.пф/anticrisis/mishustin-utverdil-perechen-naibolee-postradavshikh-ot-pandemii-otrasley-ekonomiki>

⁴<https://www.fedstat.ru/indicator/58994>

3 Эмпирический подход

Для оценки наших гипотез мы используем Survival Analysis, также известный как анализ времени до наступления события. Под событием в данном исследовании подразумевается выход отдельного индивида на рынок труда.

Первый шаг нашего подхода заключается в построении функции дожития, как наиболее упрощенной модели, описывающей эмпирическую зависимость между длительностью пребывания в статусе безработного и выходом на рынок труда. В частности, мы используем функцию выживаемости Каплана-Майера (Kaplan EL, Meier P., 1958)⁵, которую можно рассчитать следующим образом:

$$S(T) = \prod_{t=1}^T (1 - \frac{d_t}{n_t})$$

где $S(T)$ – функция дожития к моменту времени t ; t – порядковый номер времени окончания наблюдения; n_t – число индивидов, которые наблюдаются в момент времени t ; d_t – число выходов из статуса "безработного" в момент времени t ; \prod – произведение по всем наблюдениям t , завершившимся к моменту времени t .

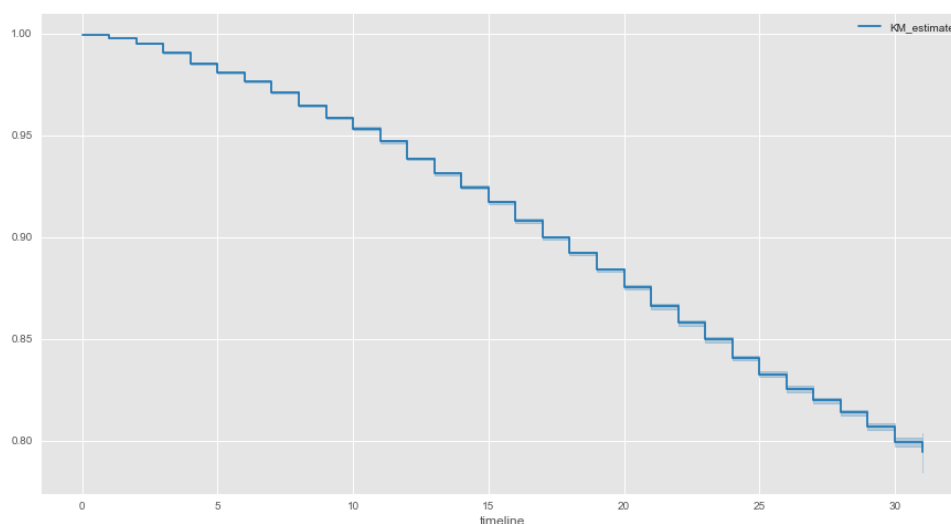


Рис. 4. Функция выживаемости Каплана-Майера

Функция Каплана-Майера показывает кумулятивную вероятность наличия статуса безработного к моменту времени t . На Рис. 4 можно видеть, что к последнему доступному периоду (32 недели) вероятность остаться безработным составляет 80%. Как уже отмечалось, этот подход основан на изменении зависимой переменной во времени и не учитывает другие факторы, влияющие на вероятность выхода на рынок.

Далее мы рассматриваем усложненную модель, учитывающую дополнительные характеристики, влияющие на выход индивида из статуса безработного. Здесь используется функция риска (hazard function) – вероятность того, что индивид найдет

⁵Kaplan E. L., Meier P. Nonparametric estimation from incomplete observations // Journal of the American statistical association. – 1958. – Т. 53. – №. 282. – С. 457-481.

работу за малый промежуток времени (выйдет из статуса безработный в статус трудоустроен). Математически функция описывается следующим образом:

$$h(T) = \lim_{\Delta T \rightarrow 0} \frac{P(T \leq t < T + \Delta T | T \leq t)}{\Delta T} = \frac{f(T)}{S(T)}$$

Кумулятивная функция риска – сумма частот индивидуальных рисков:

$$H(T) = \int_0^T h(x) dx$$

Соотношение кумулятивной функция риска с функцией дожития представлено ниже:

$$S(T) = e^{-H(T)}$$

В качестве подхода, опирающегося на функцию риска, мы используем модель пропорциональных рисков Д.Р.Кокса. В данной модели отношение между риском наступления события и независимыми регрессорами выражается следующим образом:

$$h(T) = h_0(T) * e^{\sum_{i=1}^p x_i \beta_i}$$

где x_1, \dots, x_p - независимые регрессоры, $h_0(T)$ - базовая вероятность наступления события (baseline hazard rate, одинаковый для всех наблюдений модели), β_0, \dots, β_p - коэффициенты регрессии. Данный подход использует метод частичного правдоподобия (partial likelihood), применяющийся для оценки моделей с цензурированной выборкой. В следующем разделе описаны полученные результаты.

4 Результаты

Ниже представлены итоговые результаты (таб. 2). В целом, спецификация даёт адекватные и устойчивые результаты.

Как можно увидеть из 2 все наши выдвинутые гипотезы подтвердились ⁶.

Во-первых, факт наличия детей не только не снижает вероятность трудоустройства в конкретный момент времени, но даже увеличивает на 19% при прочих равных. Мы связываем это с тем, что затраты, связанные с детьми, могли не только не сократиться с наступлением пандемии, но и даже вырасти: отсутствие бесплатных школьного питания и дополнительных занятий; необходимость обеспечить ребенка техническими средствами для дистанционного обучения; возможно, появление репетиторов, потому как качество образования особенно для школьников сильно упало в пандемию из-за дистанционного формата.

В свою очередь увеличение нетрудового дохода за счет дополнительного пособия на ребенка снижает вероятность выхода на рынок труда на 2%. Это также подтверждает одну из наших гипотез об искажении стимулов к трудоустройству.

⁶в данной модели интерпретация результатов основывается на колонках $\exp(\text{coef})$ и $p\text{-value}$. Колонка coef является "технической".

	coef	exp(coef)	lower 95%	upper 95%	p-value
covariate					
did_not_work	-0.20	0.82	0.79	0.85	0.00
social_status_has_child	0.17	1.19	1.16	1.21	0.00
cv_experience	-0.00	1.00	1.00	1.00	0.07
cv_salary	-0.01	0.99	0.99	0.99	0.00
cv_gender	-0.12	0.89	0.87	0.90	0.00
cv_retrainingcapability	0.05	1.05	1.03	1.07	0.00
sum_social_benefits	-0.02	0.98	0.98	0.98	0.00
sum_child_social_benefits	-0.03	0.98	0.97	0.98	0.00
work_2019_7_2020_9	0.95	2.59	2.51	2.68	0.00
last_salary	0.00	1.00	1.00	1.00	0.00
year_unemployment	-0.03	0.97	0.97	0.97	0.00
dam	-0.21	0.81	0.80	0.81	0.00
cv_educationtype_1	0.16	1.18	1.15	1.20	0.00
cv_educationtype_2	0.10	1.10	1.08	1.12	0.00
18_30	0.10	1.11	1.07	1.15	0.00
31_40	-0.01	0.99	0.96	1.02	0.63
41_50	0.08	1.08	1.05	1.12	0.00

Таблица 2. Оценки параметров модели пропорциональных рисков Кокса

Во-вторых, факт отсутствия трудовой деятельности снижает вероятность трудоустройства на 18%. Этот результат вполне соответствует эмпирическим закономерностям: даже в докризисные периоды многие работодатели заявляют факт наличия опыта как необходимый критерий, вполне вероятно, что в период пандемии важность этого критерия могла даже возрасти в связи с затруднениями в деловой среде.

В-третьих, подтвердилась наша гипотеза об усложнении трудоустройства в тех регионах, где большая доля экономически активного населения занята в наиболее пострадавших от коронавируса отраслях. Это может быть объяснено большим количеством простоев, ликвидаций предприятий, чем в других отраслях, следовательно, меньшим спросом на рабочую силу. Таким образом, рост этого показателя в регионе на 1% приводит к снижению вероятности трудоустройства на 19%.

И, в-четвертых, как и предполагалось, вероятность выхода на рынок труда выше всего для индивидов, имеющих высшее образование. Наличие высшего образования увеличивает вероятность трудоустройства на 18% по сравнению с группами, имеющими только среднее полное общее образование. Наличие среднего профессионального образования увеличивает вероятность трудоустройства на 10%.

На Рис. 5 и Рис. 6 можно увидеть построенные расчетные функции дожития для различных групп безработных: в зависимости от наличия статуса "отсутствие трудовой деятельности" и наличия детей. Как можно видеть, с увеличением продолжительности пребывания человека в статусе безработного, возрастает и разница

в вероятности трудоустройства в конкретный момент времени, и к 30 неделе нахождения в статусе безработного эта разница составляет около 3 п.п.

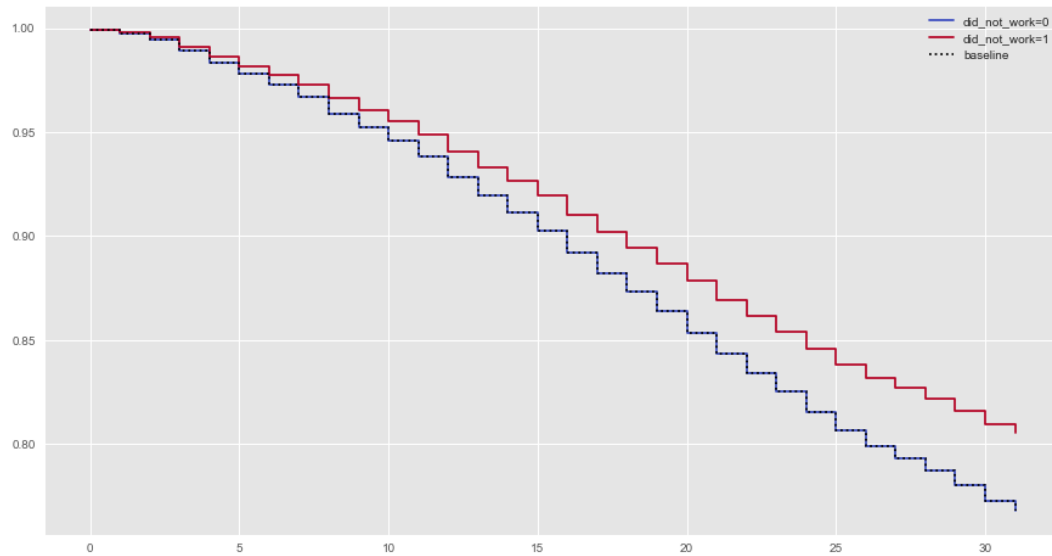


Рис. 5. Кривые дожития в зависимости от наличия статуса "отсутствие трудовой деятельности"

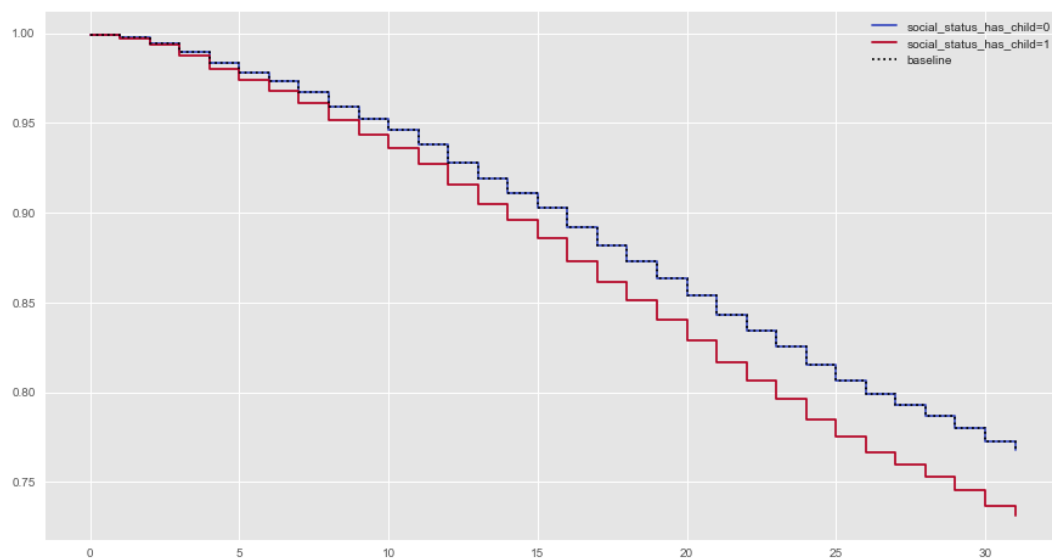


Рис. 6. Кривые дожития в зависимости от факта наличия детей