

Моделирование длительности безработицы по данным российского мониторинга экономики и здоровья¹⁾

Фурманов К.К.

Российский мониторинг экономики и здоровья (РМЭЗ) является одним из двух основных источников данных для эконометрического анализа длительности безработицы в России, наряду с данными государственных служб занятости (ГСЗ). При этом данные РМЭЗ содержат информацию о незарегистрированных безработных, которую не предоставляют ГСЗ, а дополнительная привлекательность РМЭЗ заключается в доступности широкому кругу исследователей. Однако использование данных мониторинга сопряжено со значительными трудностями: опрос индивидов проводится не чаще раза в год, что затрудняет расчет продолжительности состояний безработицы и приводит к смещению отбора. Смещение возникает, так как моменты опроса с большей вероятностью попадают на продолжительные состояния безработицы, в то время как короткие состояния чаще остаются незамеченными, «проскакивая» между опросами. В статье предлагается метод эконометрического анализа длительности безработицы, ориентированный на использование данных РМЭЗ и учитывающий их специфику. Демонстрируется применение метода к данным 2000–2005 гг. о безработице в России.

Ключевые слова: продолжительность безработицы; РМЭЗ; временная зависимость; смещение отбора.

Введение

Одной из основных тем исследований, посвященных эконометрическому анализу рынка труда, является продолжительность безработицы. Цель таких исследований заключается в выявлении факторов, определяющих вероятность выхода индивида из состояния безработицы. Кроме различных индивидуальных и региональных характеристик на эту вероятность влияет продолжительность пребывания в безработном состоянии. Зависимость вероятности выхода из некоторого состояния от продолжительности пребывания в нем называется временной зависимостью (duration или temporal dependence), и выявление характера временной зависимости

¹⁾ Автор выражает благодарность Т.А. Ратниковой, Э.Б. Ершову, Г.Г. Канторовичу, И.Г. Поспелову и Е.В. Коссовой за помощь и комментарии.

Фурманов К.К. – магистр экономики, преподаватель кафедры математической экономики и эконометрики, аспирант факультета экономики Государственного университета – Высшей школы экономики.

Статья поступила в Редакцию в августе 2009 г.

также входит в цели исследования длительности безработицы. Основополагающими в этой области стали работы Никеля [13] и Ланкастера [12].

Те исследования, в которых проводился анализ длительности безработицы в России, можно разделить на две группы в зависимости от данных, на которые опирался анализ. К первой группе относятся работы [1] и [3], в которых исследовалась длительность регистрируемой безработицы и использовались данные государственных служб занятости (ГСЗ). Работы второй группы (к ней относятся статьи [2, 8, 9]) опираются на данные Российского мониторинга экономики и здоровья (РМЭЗ) и включают в выборку как зарегистрированных, так и незарегистрированных в ГСЗ безработных. При этом вопрос о том, каким критериям должен удовлетворять индивид, чтобы считать его безработным, решается различными исследователями по-разному.

В ГСЗ регистрируется небольшая доля безработных, и неизвестно, насколько результаты, полученные на основании данных о зарегистрированных безработных, могут быть распространены на оставшееся большинство ищущих работу индивидов. Кроме того, начало периода безработицы не совпадает с регистрацией в ГСЗ, а выход из состава зарегистрированных безработных не означает прекращения поиска работы. Эти причины побуждают исследователей обращаться к данным РМЭЗ, а дополнительная привлекательность этих данных заключается в их доступности широкому кругу исследователей.

Тем не менее использование данных РМЭЗ сопряжено со значительными трудностями. Опрашиваемые индивиды наблюдаются всего раз в год (а в период 1996–2000 гг. раз в два года), вследствие чего сложно рассчитать длительность наблюдаемых периодов безработицы. Более того, возникает смещение отбора: чаще наблюдаются продолжительные состояния безработицы, в то время как короткие состояния «проскакивают» между опросами.

В предлагаемой вниманию читателя статье описывается метод анализа длительности состояний, ориентированный на использование данных последовательности опросов с продолжительными промежутками между наблюдениями. Статья развивает идеи, заложенные в работах [5] и [6]. В работе [5] было указано на смещение отбора, возникающее при анализе данных РМЭЗ, предложен и опробован метод оценивания средней продолжительности безработицы и закона распределения длительностей безработицы по данным РМЭЗ. Статья [6] дополнила статью [5] новыми свидетельствами в пользу полученных выводов о средней продолжительности безработицы, а также результатами статистического эксперимента, демонстрирующего последствия смещения отбора. Полученные результаты относились к «общей массе» безработных, без учета индивидуальных различий. В данной работе предлагается эконометрическая модель длительности безработицы, в которой оценивается связь вероятности выхода из безработицы с индивидуальными характеристиками безработных и характеристиками регионов.

Описание модели

Мы будем различать три состояния индивида: «занятость», «безработица» и «вне рабочей силы». Нас интересует продолжительность пребывания индивида в состоянии безработицы. Эта продолжительность – случайная величина, закон распределения которой определяется вероятностями выхода в занятость и из рабочей

силы. Таким образом, моделирование длительности безработицы сводится к моделированию вероятности выхода из этого состояния.

Обозначим $p_U(j)$ вероятность того, что ищущий работу индивид, пребывающий в состоянии безработицы уже j месяцев, останется безработным в течение следующего месяца. Вероятность того, что индивид, являющийся безработным в течение j месяцев, найдет работу в течение месяца, обозначим $p_E(j)$, а того, что он выйдет из состава рабочей силы – $p_N(j)$. Ясно, что $p_U(j) + p_E(j) + p_N(j) = 1$.

К сожалению, мы не можем оценить величины $p_U(j)$, $p_E(j)$ и $p_N(j)$ непосредственно из имеющихся у нас данных, так как для этого необходимо производить наблюдения за безработными в каждом месяце. Данные РМЭЗ, на которых основывается наш анализ, получены в результате опросов, проводящихся с интервалом в год, в связи с этим мы вынуждены делать некоторые предположения о том, как ведут себя интересующие нас вероятности.

Простейшее из таких предположений заключается в том, что вероятности нахождения работы и выхода из рабочей силы неизменны с течением времени. Хотя такая предпосылка и представляется весьма жесткой, полученные с ее помощью результаты можно рассматривать как первое приближение для оценки средней длительности безработицы.

Более гибким является подход, при котором мы считаем вероятности выхода из безработицы постоянными в течение некоторых временных интервалов, но меняющимися от одного интервала к другому. Например, мы можем разбить всю временную ось на интервалы 0–5 мес., 6–12 мес., более 12 мес. и считать вероятность нахождения работы неизменной внутри каждого из этих интервалов. Различные предположения о поведении вероятностей выхода, использованные нами, будут описаны подробно в дальнейшем.

Из имеющихся у нас данных мы можем определить, закончилось ли состояние безработицы, зафиксированное в некотором опросе, к моменту следующего опроса того же индивида. Пусть на момент опроса безработный искал работу уже t месяцев, а до момента следующего опроса прошло l месяцев. Тогда состояние безработицы, наблюдавшееся при первом опросе, будет продолжаться до второго опроса со следующей вероятностью $U(t, l)$:

$$(1) \quad U(t, l) = \prod_{j=t}^{t+l-1} p_U(j).$$

Теперь определим вероятность $E(t, l)$ того, что наблюдавшееся состояние закончится к моменту следующего опроса тем, что безработный найдет работу. Безработный может найти работу сразу же после первого опроса с вероятностью $e(t)$, либо в течение месяца оставаться безработным, а потом устроиться (вероятность такого случая равна $c(t) \cdot e(t+1)$), либо два месяца искать работу и найти ее на третий и т.д. Сложив вероятности этих случаев, получаем:

$$(2) \quad E(t, l) = p_E(t) + p_U(t)p_E(t+1) + p_U(t)p_U(t+1)p_E(t+2) + \dots = \sum_{j=t}^{t+l-1} \left(p_E(j) \prod_{k=t}^{j-1} p_U(k) \right).$$

Аналогично выводится выражение и для вероятности выхода из состава рабочей силы:

$$(3) \quad N(t, l) = \sum_{j=t}^{t+l-1} \left(p_N(j) \prod_{k=t}^{j-1} p_U(k) \right).$$

Отметим, что формулы (1), (2) и (3) не отражают вероятности, с которой индивид, бывший на момент первого опроса безработным, во время второго опроса будет являться безработным, занятым или экономически неактивным. Это – вероятности того, что именно то состояние безработицы, которое наблюдалось при первом опросе, закончится выходом в занятость или из рабочей силы (либо не закончится) к моменту второго опроса.

Выражения (2) и (3) будем называть уравнениями выхода в занятость и из состава рабочей силы.

Так как имеющиеся у нас данные позволяют определить, нашел ли респондент работу между опросами, прекратил ли поиск или остался безработным, то, сделав вышеупомянутые предположения о характере функций $p_U(t)$, $p_E(t)$ и $p_N(t)$, мы можем оценить эти функции с помощью метода максимального правдоподобия, что и было нами сделано. Для получения этих оценок мы использовали программный пакет Stata версии 10.0.

Распространение вышеизложенного подхода на случай анализа более чем двух опросов не составляет труда. Так, например, вероятность того, что респондент будет продолжать искать работу между первым и вторым опросами и найдет ее ко времени третьего опроса, равна $U(t, l_1) \cdot E(t+l_1, l_2)$, где l_1 и l_2 – число месяцев между первым и вторым и между вторым и третьим опросами соответственно. Рассчитывая такие вероятности, мы опять же можем применить метод максимального правдоподобия для получения интересующих нас оценок. Более того, для применения метода максимального правдоподобия неважно, имеем ли мы одно наблюдение, в котором респондент нашел работу между вторым и третьим раундами, безуспешно искав работу между первым и вторым раундами, или имеются два наблюдения, в одном из которых респондент остался безработным, а в другом – нашел работу. Таким образом, случай с большим количеством опросов сводится к случаю с двумя раундами опроса, но с большим числом наблюдений.

Вероятности выхода в различные состояния зависят, конечно, не только от времени пребывания в состоянии безработицы. На них влияют различные факторы, такие как индивидуальные характеристики безработных и региональные характеристики рынка труда. Кроме того, степень влияния этих факторов, равно как и форма временной зависимости, нам неизвестны и подлежат оцениванию. Введем следующие обозначения для величин, определяющих вероятности выхода в безработицу и экономическую неактивность:

$x_{E,i}$, $x_{N,i}$ – векторы детерминант вероятностей выходов в различные состояния, переменные, описывающие различия между ищущими работу индивидами. Индекс i означает номер соответствующего индивида;

β_E , β_N – коэффициенты при соответствующих переменных из векторов $x_{E,i}$, $x_{N,i}$, описывающие направление и степень влияния детерминант на вероятности выхода;

α_E , α_N – параметры, описывающие степень и знак временной зависимости.

Как обычно в эконометрике, будем считать, что влияние объясняющих переменных сводится к влиянию их линейных комбинаций с коэффициентами: $x'_{E,i}\beta_E$ и $x'_{N,i}\beta_N$ (апостроф здесь обозначает транспонирование). Что касается временной зависимости, то она описывается некими функциями, зависящими от параметров α_E и α_N . Обозначим эти функции $\varphi_E(t | \alpha_E)$ и $\varphi_N(t | \alpha_N)$. В качестве φ_E и φ_N могут быть выбраны, например, кусочно-постоянные или полиномиальные функции.

Зависимость вероятностей выхода от вышеперечисленных параметров может быть задана различным образом. Мы будем использовать спецификацию, основанную на логистической модели множественного выбора (multinomial logit model):

$$(4) \quad \begin{aligned} p_E(j | x_{E,i}, x_{N,i}, \beta_E, \beta_N, \alpha_E, \alpha_N) &= \\ &= \frac{\exp(x'_{E,i}\beta_E + \varphi_E(j | \alpha_E))}{1 + \exp(x'_{E,i}\beta_E + \varphi_E(j | \alpha_E)) + \exp(x'_{N,i}\beta_N + \varphi_N(j | \alpha_N))}, \end{aligned}$$

$$(5) \quad \begin{aligned} p_N(j | x_{E,i}, x_{N,i}, \beta_E, \beta_N, \alpha_E, \alpha_N) &= \\ &= \frac{\exp(x'_{N,i}\beta_N + \varphi_N(j | \alpha_N))}{1 + \exp(x'_{E,i}\beta_E + \varphi_E(j | \alpha_E)) + \exp(x'_{N,i}\beta_N + \varphi_N(j | \alpha_N))}, \end{aligned}$$

$$(6) \quad \begin{aligned} p_U(j | x_{E,i}, x_{N,i}, \beta_E, \beta_N, \alpha_E, \alpha_N) &= \\ &= \frac{1}{1 + \exp(x'_{E,i}\beta_E + \varphi_E(j | \alpha_E)) + \exp(x'_{N,i}\beta_N + \varphi_N(j | \alpha_N))}. \end{aligned}$$

Такая спецификация дает возможность интерпретации коэффициентов и гарантирует, что полученные вероятности действительно будут лежать в пределах от нуля до единицы, а сумма $p_U(t)$, $p_E(t)$ и $p_N(t)$ будет равна единице.

В дальнейшем будут описаны результаты оценивания с использованием следующих спецификаций временной зависимости (в списке опущены индексы E и N , так как спецификация для разных типов выхода задавалась одним видом функции, но с разными параметрами):

- постоянная: $\varphi(j) = 0$;
- квадратичная: $\varphi(j | \alpha) = \alpha_1 j + \alpha_2 j^2$;
- кусочно-постоянная: $\varphi(j | \alpha) = \begin{cases} 0, & j < t_1, \\ \alpha_1, & t_1 \leq j < t_2, \\ \dots \\ \alpha_p, & j \geq t_p; \end{cases}$

$$\bullet \text{ кусочно-линейная: } \varphi(j | \alpha) = \begin{cases} \frac{j}{t_1} \alpha_1, & j < t_1, \\ \alpha_1 + \frac{(j-t_1)}{(t_2-t_1)} (\alpha_2 - \alpha_1), & t_1 \leq j < t_2, \\ \dots \\ \alpha_{p-1} + \frac{(j-t_{p-1})}{(t_p-t_{p-1})} (\alpha_p - \alpha_{p-1}) & t_{p-1} \leq j < t_p, \\ \alpha_p & t \geq t_p. \end{cases}$$

Значения t_1, t_2, \dots, t_p (границы отрезков для кусочно-заданных функций) не оценивались, а указывались заранее. Во всех спецификациях нет свободного члена, $\varphi(0) = 0$, так как свободный член включался в вектор объясняющих переменных.

Также использовались смешанные спецификации: поскольку для некоторых индивидов по данным РМЭЗ невозможно рассчитать время начала безработицы (в частности, для неимеющих опыта работы), то для них предполагалось отсутствие временной зависимости (нулевые функции φ_E и φ_N), в то время как для всех остальных временная зависимость учитывалась с помощью какой-либо из других спецификаций. Кроме того, в модель вводилась отдельная переменная, отличавшая индивидов, не указавших время начала безработицы.

Задача оценивания параметров модели (векторов $\beta_E, \beta_N, \alpha_E, \alpha_N$) решается относительно просто. Обозначим S_U множество наблюдений, в которых состояние безработицы не закончилось между раундами соседних опросов. Аналогично, S_E будет множеством наблюдений, в которых состояние безработицы завершилось выходом в занятость, а S_N – множеством наблюдений, в которых произошел выход из состава рабочей силы.

Зная с точностью до параметров вероятности p_U, p_E и p_N , можно рассчитать вероятности попадания наблюдения в каждое из множеств, т.е. величины U, E и N , задаваемые формулами (1)–(3) и, естественно, зависящие от наблюдаемых характеристик индивидов $x_{E,i}, x_{N,i}$, продолжительности поиска работы на момент первого опроса t_i , длины временного отрезка между двумя раундами опроса l_i и от параметров модели $\beta_E, \beta_N, \alpha_E, \alpha_N$. Задача оценивания параметров сводится к задаче максимизации логарифмической функции правдоподобия:

$$\begin{aligned} & \ln L(x_{E,1}, \dots, x_{E,n}; x_{N,1}, \dots, x_{N,n}; t_1, \dots, t_n; l_1, \dots, l_n; \beta_E, \beta_N, \alpha_E, \alpha_N) = \\ (7) \quad & = \sum_{i \in S_U} \ln U(t_i, l_i | x_{E,i}, x_{N,i}; \beta_E, \beta_N, \alpha_E, \alpha_N) + \sum_{i \in S_E} \ln E(t_i, l_i | x_{E,i}, x_{N,i}; \beta_E, \beta_N, \alpha_E, \alpha_N) + \\ & + \sum_{i \in S_N} \ln N(t_i, l_i | x_{E,i}, x_{N,i}; \beta_E, \beta_N, \alpha_E, \alpha_N) \rightarrow \max_{\beta_E, \beta_N, \alpha_E, \alpha_N}, \end{aligned}$$

где n – число наблюдений.

Максимизация проводилась с помощью метода Ньютона – Рафсона при использовании пакета Stata версии 7.0. Никаких проблем при максимизации обнару-

жено не было, хотя в случае использования некоторых других спецификаций временной зависимости (в данной работе не применявшихся) возникали множественные максимумы.

Ненаблюдаемая разнородность

Несмотря на включение объясняющих переменных в модель, могут оставаться неучтенные факторы, влияющие на вероятности выхода. То остаточное различие, которое не удалось учесть объясняющими переменными, называется ненаблюдаемой разнородностью (*unobserved heterogeneity*).

Попытки учесть ненаблюдаемую разнородность часто встречаются в анализе панельных данных, где для ее моделирования используются модели с детерминированным или случайным индивидуальным эффектом (см., например, [4]). Наиболее распространенные в анализе длительностей подходы можно рассматривать как разнородности модели со случайным эффектом. При таком подходе суммарное воздействие ненаблюдаемых факторов моделируется с помощью случайной величины, про которую обычно предполагается, что она является независимой от объясняющих переменных. Влияние ненаблюдаемых параметров на вероятность выхода складывается с влиянием объясняющих переменных – их общее влияние можно описать выражением типа $(x'\beta + v)$, где v – случайный индивидуальный эффект.

Отметим, что в нашем случае предположение о независимости является неправдоподобным. Дело в том, что среди экзогенных переменных в модели присутствует продолжительность пребывания в состоянии безработицы t (хотя мы и рассматриваем эту переменную отдельно от векторов x_N и x_E). Скорее всего, среди индивидов, ищущих работу продолжительное время, доля людей с меньшей вероятностью нахождения работы выше, чем среди только что вошедших в состояние безработицы, а значит, величина t коррелирует с индивидуальным эффектом. Эта проблема получила название проблемы начальных условий (*initial conditions problem*, см. [10]), и ее решение весьма затруднительно. Рассмотрим только простой случай, в котором временная зависимость отсутствует и величина t не влияет на вероятность выхода.

Пусть ненаблюдаемая разнородность описывается моделью разделенной совокупности (*split population model*, см. [14]), называемой также моделью *mover-stayer*. В такой модели предполагается наличие двух латентных классов. Представители одного из этих классов имеют шансы найти работу, а представители другого обречены на неудачу. Таким образом, случайный эффект принимает два возможных значения: ноль и минус бесконечность. Данный подход к моделированию длительности безработицы был использован, например, в работе Фолея [8].

Причиной для такого различия может являться то, что люди при опросе указывают, что они ищут работу, однако в самом деле никаких попыток поиска не предпринимают. Будем считать, что ненаблюдаемое различие влияет только на вероятность нахождения работы – каждый человек в состоянии выйти из состава рабочей силы. Вероятность принадлежности индивида к классу «безнадежных» обозначим p .

Функция правдоподобия в данном случае несколько усложняется из-за того, что мы больше не считаем имеющуюся выборку независимой, так как зависимы между собой наблюдения, относящиеся к одному индивиду. Величины, меняющиеся

от наблюдения к наблюдению, будем записывать с парой индексов i (номер индивида) и m (номер наблюдения за индивидом). Случайный эффект v меняется от индивида к индивиду но постоянен между наблюдениями, так что его запись будет сопровождаться одним индексом v_i .

Для «безнадежных» значение случайного эффекта v_i равно минус бесконечности, а значит, $\exp(x'_{E,i}\beta + v_i) = 0$. В связи с этим вероятности p_U , p_E и p_N при условии принадлежности к данному классу выглядят следующим образом:

$$(8) \quad p_E(j | x_{E,i,m}, x_{N,i,m}, \beta_E, \beta_N, v_i = -\infty) = 0,$$

$$(9) \quad p_N(j | x_{E,i,m}, x_{N,i,m}, \beta_E, \beta_N, v_i = -\infty) = \frac{\exp(x'_{N,i,m}\beta_N)}{1 + \exp(x'_{N,i,m}\beta_N)},$$

$$(10) \quad p_U(j | x_{E,i,m}, x_{N,i,m}, \beta_E, \beta_N, v_i = -\infty) = \frac{1}{1 + \exp(x'_{N,i,m}\beta_N)}.$$

Обратим внимание на то, что в записи отсутствуют параметры α_E , α_N и функции ϕ_E и ϕ_N , так как мы описываем модель без временной зависимости. Для класса безработных, имеющих шансы на трудоустройство ($v_i = 0$), вероятности выхода в течение месяца, как и раньше, описываются выражениями (4)–(6).

Аналогичную поправку внесем в выражения для вероятностей выхода в течение времени между раундами опроса. Для класса «безнадежных» вероятности $U(t, l)$, $E(t, l)$ и $N(t, l)$ приобретают следующие выражения:

$$(11) \quad \begin{aligned} & U(t_{i,m}, l_{i,m} | x_{E,i,m}, x_{N,i,m}, \beta_E, \beta_N, v_i = -\infty) = \\ & = \prod_{j=t}^{t+l-1} p_U(j | x_{E,i,m}, x_{N,i,m}, \beta_E, \beta_N, v_i = -\infty), \end{aligned}$$

$$(12) \quad E(t_{i,m}, l_{i,m} | x_{E,i,m}, x_{N,i,m}, \beta_E, \beta_N, v_i = -\infty) = 0,$$

$$(13) \quad \begin{aligned} & N(t_{i,m}, l_{i,m} | x_{E,i,m}, x_{N,i,m}, \beta_E, \beta_N, v_i = -\infty) = \\ & = 1 - U(t_{i,m}, l_{i,m} | x_{E,i,m}, x_{N,i,m}, \beta_E, \beta_N, v_i = -\infty). \end{aligned}$$

Опустим выражение для функции правдоподобия из-за его громоздкости. По сути, единственное усложнение по сравнению с выражением (7) связано с наличием зависимости между наблюдениями. К оцениваемым параметрам модели, по которым максимизируется функция правдоподобия, добавляется вероятность принадлежности к классу «безнадежных» p . Тем не менее, в связи с уже упомянутыми причинами, из модели исключаются параметры α_E и α_N .

Интерпретация коэффициентов

Интерпретация коэффициентов во многих нелинейных моделях затруднительна, и, к сожалению, предложенная модель не является исключением. Часто для трак-

товки полученных оценок приходится рассчитывать предельные эффекты, которые не являются универсальными – в том смысле, что предельный эффект от изменения некоторой объясняющей переменной может зависеть от значений этой и других переменных. Тем не менее определение вероятностей выхода через множественную логистическую связку позволяет дать коэффициентам интерпретацию – не слишком простую, однако верную для любого объекта вне зависимости от его индивидуальных характеристик.

Обратим внимание на отношение вероятности выхода, например, в состояние занятости к вероятности остаться в категории безработных (будем называть частное этих вероятностей *отношением выбытия*²⁾ в состоянии занятости). Из формул (4) и (6) видно, что это отношение равно

$$(14) \quad \frac{p_E(j | x_{E,i}, x_{N,i}, \beta_E, \beta_N, \alpha_E, \alpha_N)}{p_U(j | x_{E,i}, x_{N,i}, \beta_E, \beta_N, \alpha_E, \alpha_N)} = \exp(x'_{E,i} \beta_E + \varphi_E(j | \alpha_E)).$$

Обозначим β^* коэффициент при некоторой переменной x^* из вектора $x_{E,i}$. Если величина x^* вырастет на единицу, то отношение выбытия вырастет в $\exp(\beta^*)$ раз, т.е.

$$(15) \quad \frac{\exp(x'_{E,i} \beta_E + \beta^* + \varphi_E(j | \alpha_E))}{\exp(x'_{E,i} \beta_E + \varphi_E(j | \alpha_E))} = \exp(\beta^*).$$

Таким образом, изменение значения объясняющей переменной приводит к пропорциональному изменению отношения выбытия с коэффициентом пропорциональности, равным потенцированному значению коэффициента при изменяемой переменной. То же самое можно выразить иначе: при увеличении x^* на единицу отношение выбытия вырастет на $(\exp(\beta^*) - 1) \times 100\%$.

Аналогичные выводы можно сделать и для отношения выбытия из рабочей силы.

К сожалению, рост отношения выбытия в некоторое состояние не означает увеличения вероятности выхода в это состояние (см., например, [7, с. 46–47]). Однако можно показать, что при увеличении отношения выбытия растет условная вероятность выхода в соответствующее состояние при условии, что выхода в другие состояния не произойдет. Это следует из того, что указанная условная вероятность описывается обычной логит-моделью (как в формуле (9), отражающей вероятность выхода из рабочей силы для индивида, про которого точно известно, что на работу он не устроится).

Данные

Для анализа использовались данные РМЭЗ за 2000–2005 гг. (IX–XVI раунды опроса). Опрос домохозяйств проводился в конце каждого года анализируемого периода.

²⁾ Для этого показателя есть краткий англоязычный термин «odds», однако столь же краткого русского аналога мне найти не удалось.

При составлении выборки безработных использовался тот же подход, что и в работе [5]. Основой для определения индивида как безработного являлось определение безработицы, данное Международной организацией охраны труда (МОТ). Согласно этому определению, безработным считается индивид:

- не имеющий работы;
- желающий найти ее и активно ищущий;
- готовый приступить к работе в ближайшее время.

Однако были введены дополнительные ограничения. Так, в анализируемую подвыборку безработных попали только индивиды в возрасте от 18 до 60 лет включительно независимо от пола. Хотя российская служба статистики включает в состав безработных индивидов от 15 до 72 лет, крайние возрастные группы были исключены, так как для них закон распределения длительности безработицы может сильно отличаться. По тем же причинам в выборку не включались индивиды, трудоустройству которых мешает плохое состояние здоровья.

Для того чтобы определить, был ли осуществлен выход из безработицы между двумя раундами опроса, необходимо, чтобы индивид наблюдался в обоих раундах. Вследствие этого возникало истощение выборки, так как некоторые индивиды выбывали из выборки РМЭЗ. В результате итоговая подвыборка состояла из 2010 наблюдений за 1507 индивидами. Отметим, что так как для отслеживания переходов между состояниями нужны данные двух опросов, то данные, соответствующие одному индивиду в двух раундах, составляют одно наблюдение, в трех раундах – два наблюдения и т.д.

Если безработный в n -ом раунде респондент в следующем раунде опроса указывал, что у него есть работа, считалось, что из безработицы он вышел в занятость. Тот же вывод делался, если респондент указывал, что с последней работы он ушел после раунда n , пусть даже в раунде $n + 1$ он опять был безработным (т.е. за промежуток времени между раундами он перешел из безработицы в занятость, а потом обратно). Респондент, не нашедший работу между раундами n и $n + 1$, но и не оставшийся в составе безработных, классифицировался как вышедший из рабочей силы. Если в течение периода безработицы индивид стал старше 60 лет, учитывались только те наблюдения за ним, в которых его возраст удовлетворял нашим ограничениям, после 60 лет этот респондент считался выбывшим из рассмотрения (но не вышедшим из рабочей силы). Аналогично, не рассматривались наблюдения за индивидами младше 18 лет, но как только эти индивиды достигали совершеннолетия, они могли классифицироваться как безработные.

Более подробное описание составления анализируемой подвыборки можно найти в работе [5].

В выборе объясняющих переменных мы опирались на работы Фолея и Гроган и ван ден Берга. Данные об индивидуальных характеристиках безработных и месте их проживания брались из базы РМЭЗ, данные о региональном уровне безработицы – из сборников Федеральной службы государственной статистики РФ. Так как опросы РМЭЗ проводились в конце года, то промежутку между раундами, например 2000 и 2001 гг., в соответствие ставился уровень безработицы 2001 г. Значения объясняющих переменных считались не изменяющимися в промежутке между раундами (кроме, конечно, времени пребывания в состоянии безработицы).

Единственное значительное отличие от работ Фолея [8], Гроган и ван ден Берга [9] и Карцевой [2] в выборе объясняющих переменных заключается в учете дохо-

дов и расходов индивидов и домохозяйств. Гроган и ван ден Берг использовали данные о заработной плате в один из предыдущих состояниях безработицы раундов и дамми-переменную для тех, чья заработная плата была неизвестна. Фолей включал в модель расходы домохозяйств. В работе Карцевой эти факторы в модель не включались.

Хотя расходы домохозяйства представляются важным фактором, влияющим на вероятность нахождения работы, его использование сопряжено с рядом трудностей. Во-первых, расходы домохозяйств нелегко рассчитать, и мы не знаем, насколько наши расчеты будут точны. Во-вторых, расходы стоит представить в расчете на одного человека, и здесь встает вопрос, какую нагрузку создает каждый из членов домохозяйства. В-третьих, данные о расходах имеются в виде ответов на множество вопросов, и далеко не каждое домохозяйство отвечает на каждый из них. Таким образом, возникает истощение выборки.

В связи с вышеперечисленными проблемами в настоящем исследовании вместо данных о расходах и доходах использовалась субъективная оценка богатства, приведенная индивидом при ответе на следующий вопрос (№ 62 в опроснике РМЭЗ для взрослых):

«Представьте себе, пожалуйста, лестницу из 9 ступеней, где на нижней, первой, ступени стоят нищие, а на высшей, девятой – богатые. На какой ступени находитесь сегодня вы лично?».

На рис. 1 приведена гистограмма распределения ответов на этот вопрос в анализируемой подвыборке. За базовую категорию принимались индивиды, ставившие себя на третью, четвертую и пятую ступени. Относительно них с помощью дамми-переменных выделялись две категории: бедных и богатых. Отметим, что на графике отображена доля ответов среди наблюдений, а не среди респондентов – одному респонденту могло соответствовать несколько наблюдений.

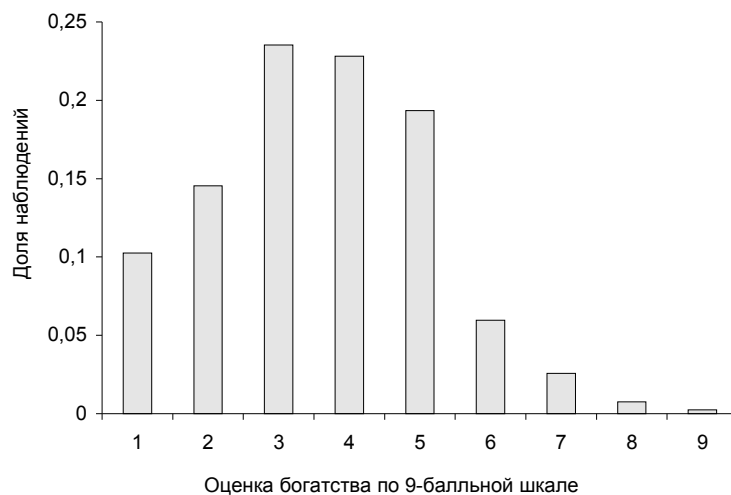


Рис. 1. Распределение наблюдений анализируемой подвыборки по субъективной оценке богатства респондентов

Оговорим также определение дамми-переменной, отделяющей безработных, имеющих опыт работы, от неопытных. Считалось, что индивид имеет опыт работы, если он давал положительный ответ на соответствующий вопрос. Затруднявшиеся ответить приравнивались к неопытным.

В табл. 1 приведено распределение наблюдений анализируемой подвыборки по объясняющим качественным переменным, использованным в модели. Единственная количественная переменная – региональный уровень безработицы – имеет среднее значение 9,4% и стандартное отклонение 4,9%. Минимальный уровень безработицы наблюдался в Москве в 2005 г. (0,8%), максимальный – в Кабардино-Балкарии в 2004 г. (25,7%). Соответствующая переменная измерялась в процентах (т.е. при уровне безработицы в 5% переменная была равна 5).

Таблица 1.
Состав анализируемой подвыборки по объясняющим переменным

Объясняющие качественные переменные	В процентах
Женщины	49,4
Мужчины	50,6
Возраст:	
от 50 лет (включительно)	11,5
25-49 лет	60,5
до 24 лет (включительно)	28,0
Высшее образование	11,6
Образование ниже среднего	20,0
Отсутствие опыта работы	12,5
Зарегистрированы в государственной службе занятости	11,8
Состоят в браке	46,6
Относят себя к категории:	
бедных (оценки 1 и 2 по девятибалльной шкале)	24,4
богатых (оценки 6-9 по девятибалльной шкале)	9,4
Проживают в:	
Москве или Санкт-Петербурге	9,3
областном центре (но не в Москве или Санкт-Петербурге)	26,8
городе (но не областном центре)	25,8
поселке городского типа	9,7
селе	28,4
Наблюдения, относящиеся к интервалу:	
2000-2001 гг. (IX-X раунды)	21,0
2001-2002 гг. (X-XI раунды)	20,8
2002-2003 гг. (XI-XII раунды)	20,2
2003-2004 гг. (XII-XIII раунды)	19,4
2004-2005 гг. (XIII-XIV раунды)	18,6

Результаты

По описанным выше данным оценивались пять основных моделей длительности безработицы, отличающиеся различными предположениями о функциональной спецификации временной зависимости и наличии ненаблюдаемой разнородности:

- модель с отсутствием временной зависимости;
- модель с отсутствием временной зависимости, но наличием ненаблюдаемой разнородности типа *mover-stayer*;
- модель с временной зависимостью, описываемой квадратичными функциями φ_E и φ_N ;
- модель с временной зависимостью, описываемой кусочно-постоянными функциями φ_E и φ_N ;
- модель с временной зависимостью, описываемой кусочно-линейными функциями φ_E и φ_N .

Во всех моделях, кроме второй, ненаблюдаемая неоднородность не учитывалась по уже указанной ранее причине возникновения проблемы начальных условий. В кусочно-заданных моделях использовались следующие интервалы для задания временной зависимости: 0–2 месяца, 3–11 месяцев, 12–23 месяца, 24–35 месяцев, 36 и более месяцев.

Для максимизации функции правдоподобия использовался максимизатор пакета Stata 10.0 и метод Ньютона – Рафсона. Расчет первых и вторых производных функции правдоподобия производился численно (метод *d0* для модели *mover-stayer* и метод *lf* для остальных спецификаций). При использовании кусочно-заданных функций временной зависимости была обнаружена проблемная зона невогнутости функции правдоподобия – впрочем, эта проблема оказалась легко решаемой указанием опции максимизации «difficult». В области найденных максимумов невогнутость не была обнаружена ни для одной из спецификаций.

Все оцененные модели оказываются значимыми в целом, с точным уровнем значимости (*p*-значением), практически неотличимым от нуля.

Результаты оценивания коэффициентов приведены в Приложении 1. Наряду с оценками коэффициентов приводятся и потенцированные значения, так как предложенная выше интерпретация касается именно потенцированных коэффициентов. Однако мы приводим и «обычные» значения коэффициентов (до потенцирования) – в частности, потому что они нагляднее демонстрируют направление влияния детерминант на отношение выбытия, так как могут принимать и положительные, и отрицательные значения. Также анализ непотенцированных коэффициентов более нагляден для различения детерминант по степени воздействия. Значения коэффициентов, равные, например, единице и минус единице, свидетельствуют об одинаковой степени, но разным направлении воздействия детерминант на отношение выбытия. Однако до этого вывода не так просто додуматься на основании анализа потенцированных коэффициентов, которые при этом равны соответственно 2,72 и 0,37.

В Приложении 2 также приводятся значения логарифмической функции правдоподобия и информационного критерия Акаике (AIC). Критерий Акаике широко используется для сравнения невложенных статистических моделей, оцениваемых методом максимального правдоподобия, и опирается на значение функции правдоподобия с учетом штрафа на количество оцениваемых параметров: $AIC = -2 \ln L + 2p$, где

p – количество параметров модели³⁾. Наименьшее значение критерия соответствует наилучшей модели. Согласно Хилбе [11, с. 260], различия в критерии, по абсолютному значению меньше 2,5, свидетельствуют о том, что модели обеспечивают практически одинаковое качество подгонки, но Хилбе отмечает, что строгого статистического критерия для выбора модели с помощью АИС пока нет.

Как видно, из оцененных моделей лучше всего данные описывает кусочно-постоянная, немного хуже – кусочно-линейная и квадратичная. Модели без временной зависимости обеспечивают значительно худшее качество подгонки, причем модель разделенной совокупности несколько лучше, чем «простая» модель. Однако причиной этого превосходства может являться неучтенная временная зависимость – ее довольно сложно отличить от ненаблюдаемой разнородности. Учет разнородности может улучшить подгонку именно за счет того, что она «берет на себя» различия в вероятности выхода, объясняемые разной продолжительностью пребывания в безработице.

При оценивании модели разделенной совокупности доля индивидов с нулевой вероятностью нахождения работы была оценена как 4,4%, хотя эта оценка также может быть смещена вследствие неучтенной временной зависимости.

Оценки коэффициентов и стандартных отклонений близки во всех моделях, хотя особенно сильно это сходство проявляется среди моделей с учетом временной зависимости. Можно заметить различие в оценках для свободного члена (константы) и коэффициента при дамми-переменной, выделяющей группу безработных, для которых неизвестна дата ухода с последнего места работы. Это различие связано с тем, что разные спецификации по-разному оценивают вероятности выхода в начале периода безработицы.

Из Приложения 1 видно, что в модели присутствовали переменные для учета временного эффекта. Было обнаружено, что вероятность выхода из состава рабочей силы в 2003 и 2005 гг. была относительно большей. Изначально модель предусматривала временной эффект каждого года на вероятности обоих типов выхода. При проверке значимости с помощью теста Вальда эффект оказался значимым в совокупности на уровне 5% при всех спецификациях. Однако все переменные кроме оставшихся двух оказались незначимыми на уровне 10% как по отдельности, так и в совокупности. Для экономии места приведены результаты оценивания сокращенных регрессий, которые, впрочем, практически не отличаются от оценок, полученных с использованием более полных моделей.

Отметим те выводы относительно влияния объясняющих переменных, которые оказываются устойчивыми к изменению спецификации модели. Будем считать влияние некоторого фактора значимым, только если оно оценивается как значимое во всех спецификациях. Описывая влияние тех или иных характеристик на вероятности выхода в одно из состояний, будем подразумевать условную вероятность выхода при том, что выхода в другое состояние не произойдет (см. выше параграф «Интерпретация коэффициентов»).

Пол и возраст. Индивиды старшей возрастной группы (от 50 лет) находят работу с меньшей относительно базовой категории (25–49 лет) вероятностью, а инди-

³⁾ Распространено и другое определение критерия Акаике, в котором присутствует нормировка на число наблюдений n : $AIC = -(2 \ln L + 2p)/n$, но наличие этой нормировки в данном случае несущественно.

виды младшей группы (до 24 лет) – с большей. Впрочем, самым молодым безработным найти работу может помешать отсутствие опыта.

Представители старшей возрастной группы более склонны выходить из состава рабочей силы (напомним, что выход из рабочей силы по возрасту не рассматривался – если индивид в течение периода безработицы становился старше 60 лет, он считался выбывшим из рассмотрения, а не вышедшим из безработицы).

Значимых различий между полами не обнаружено. Заранее отметим, однако, что это касается только индивидов, имеющих опыт работы.

Образование. Значимого влияния образования на вероятность выхода из безработицы не обнаружено.

Положение на рынке труда. Незначимость коэффициентов при переменной-индикаторе отсутствия опыта работы не означает, что наличие опыта работы не влияет на вероятность прекращения безработицы. Обратим внимание на то, что в модели присутствует дополнительная переменная – «не указано время увольнения», – добавленная для того, чтобы можно было включить в выборку индивидов без опыта работы, а также тех, кто не указал дату увольнения с последнего места работы. Таким образом, коэффициент при переменной опыта отражает различие между теми, у кого нет опыта работы, и теми, у кого опыт есть, но они не указали дату последнего увольнения, скорее всего, это – ушедшие с последней работы довольно давно. Полученный результат может свидетельствовать о том, что мужчина, давно не работавший, занимает на рынке труда такую же позицию, как и вообще не имеющий опыта работы.

По абсолютному значению коэффициент при переменной «не указано время увольнения» является наибольшим из всех коэффициентов в уравнении. Из этого можно сделать вывод, что наличие или отсутствие опыта работы – фактор, наиболее сильно влияющий на длительность безработицы. Что касается женщин, то на них отсутствие опыта работы сказывается в несколько меньшей степени, чем на мужчинах, о чем говорит значимый положительный коэффициент при произведении переменных пола и отсутствия опыта.

Регистрация в государственной службе занятости не увеличивает шансы найти работу, но зарегистрированные безработные меньше склонны покидать состав рабочей силы.

Семейное положение. Замужние женщины более склонны выходить из рабочей силы. Вывод довольно очевидный, но выявляется только при высоком уровне значимости (10%).

Субъективная оценка богатства. Индивиды, относящие себя к категории бедных, находят работу с меньшей вероятностью, чем остальные. Это может быть связано с тем, что у них меньше связей, которые помогли бы в трудоустройстве. Другая возможная причина заключается в том, что люди оценивают собственное благосостояние с учетом шансов найти работу, и принадлежность к категории бедных отражает некоторые ненаблюдаемые нами характеристики, отрицательно влияющие на вероятность трудоустройства. Также бедность может оказывать подавляющие и демотивирующее воздействие, приводя безработных в отчаяние.

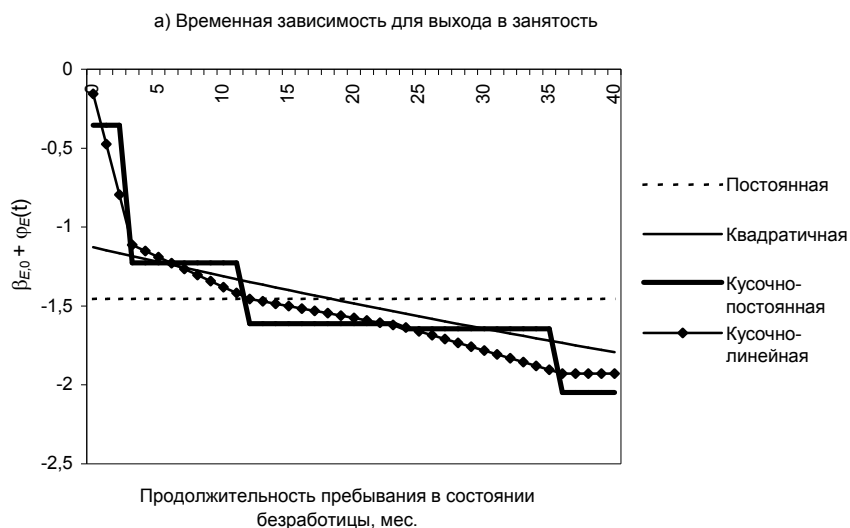
Географические характеристики. Наименьшую вероятность выхода в занятость имеют жители поселков городского типа (ПГТ). Из оценок коэффициентов видно, что проживание в ПГТ – фактор, уступающий только наличию опыта работы по силе воздействия на вероятность трудоустройства. Сельские жители проигрыва-

ют в этом отношении жителям областных центров, но заметно выигрывают у жителей ПГТ. Однако в селе также выше вероятность выхода из рабочей силы.

В регионах с более высоким уровнем безработицы меньше вероятность нахождения работы и больше вероятность выхода из состава рабочей силы.

Отрицательный знак коэффициента при индикаторе проживания в Москве и Санкт-Петербурге в уравнении выхода в занятость вовсе не означает, что для этих городов свойственна более продолжительная безработица. Это связано, во-первых, с тем, что в Москве и Санкт-Петербурге низкий уровень безработицы. Отрицательный знак может говорить о том, что продолжительность безработицы в Москве и Санкт-Петербурге больше, чем можно было ожидать от городов с таким уровнем безработицы. Но есть и более веская причина. Переменная уровня безработицы фиксировала средний уровень безработицы в регионе – как в центре, так и на периферии. Различия внутри региона учитывались с помощью дамми-переменных, определяющих тип поселения. Однако для указанных двух городов уровень безработицы рассчитывается отдельно – он относится именно к Москве или к Санкт-Петербургу. Таким образом, он уже учитывает превосходство областных центров над другими поселениями.

Временная зависимость. На рис. 2 представлены оценки функций φ_E и φ_N , отражающих временную зависимость, для разных спецификаций. Для наглядности в приведенных графиках к функциям φ_E и φ_N прибавлялись оценки свободных членов $\beta_{E,0}$ и $\beta_{N,0}$ для уравнений выхода в незанятость и из рабочей силы. Это было сделано для того, чтобы графики не только показывали характер временной зависимости, но и позволяли сравнивать отношение выбытия в разных спецификациях.



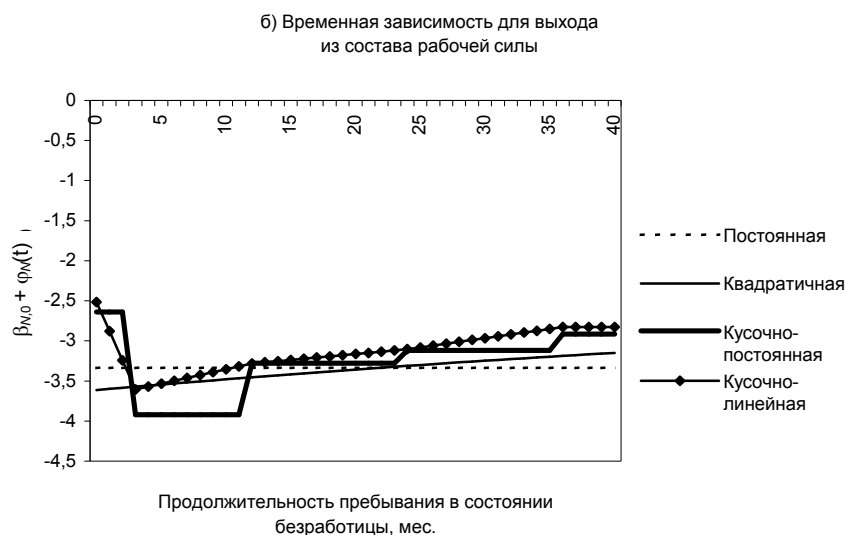


Рис. 2. Оценки временной зависимости для различных спецификаций φ_E и φ_N

Как видно, вероятность нахождения работы со временем падает, в то время как вероятность выхода из рабочей силы ведет себя немонотонно – резко падает вначале, после чего возрастает. Квадратичная спецификация не обнаруживает этой немонотонности в силу недостаточной гибкости. Она также оказывается не в состоянии описать значительное падение вероятности нахождения работы, происходящее на первых месяцах безработицы. Тот факт, что кусочно-заданные модели позволяют выявить резкие колебания вероятности выхода, приводит к различию оценок для свободного члена и коэффициента при дамми-переменной «не указано время увольнения» между кусочно-заданными и остальными моделями.

Для большей наглядности можно обратиться к рис. 3, где отражены оценки вероятности нахождения работы для некоторых типов безработных. На этом рисунке приведены графики, показывающие, как меняется вероятность нахождения работы в течение ближайшего месяца по мере роста продолжительности пребывания в состоянии безработицы. Говоря о вероятности, мы снимаем ранее сделанную оговорку о том, что выхода в другое состояние не произойдет. Приведенный график отражает вероятность нахождения работы с учетом того, что может произойти выход из состава рабочей силы.

График отражает различия в вероятности трудоустройства в зависимости от пола и наличия опыта работы. Остальные характеристики при расчете предполагались следующими:

- индивиды живут в областном центре, но не в Москве или Санкт-Петербурге;
- уровень безработицы в регионе проживания равен 5%;
- индивиды имеют среднее, но не имеют высшего образования;
- принадлежат к средней возрастной категории (25–49 лет);
- не состоят в браке;
- относят себя к средней категории при оценке собственного богатства;

- временной эффект отсутствует (т.е. наблюдается закономерность, выявленная для 2001, 2002 и 2004 гг.).

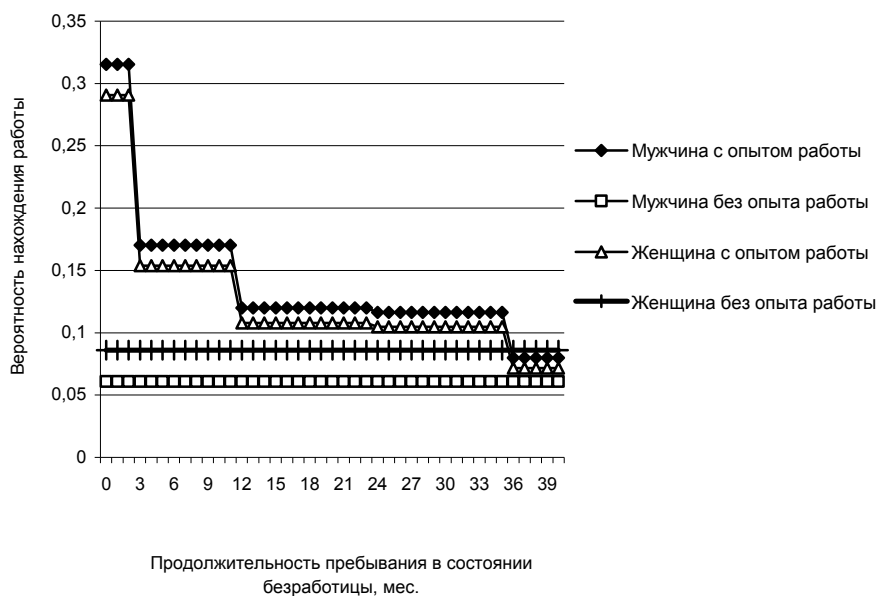


Рис. 3. Вероятность нахождения работы в течение месяца в зависимости от пола, наличия опыта работы и времени пребывания в состоянии безработицы

Линии, соответствующие индивидам без опыта работы, горизонтальны, так как для таких индивидов мы не можем установить начало периода безработицы и учесть временную зависимость. Как видно, после трех лет отсутствия работы вероятность трудоустройства приближается к уровню, характерному для индивидов, вообще не имеющих трудового опыта. Таким образом, временная зависимость, наряду с наличием опыта работы, является одним из основных факторов, влияющих на вероятность нахождения работы.

Проверка спецификации. Для проверки спецификации была осуществлена процедура, аналогичная тесту Рамсея. Исходя из полученных оценок коэффициентов рассчитывались величины $\left(x'_{E,i} \hat{\beta}_E\right)^2$, $\left(x'_{E,i} \hat{\beta}_E\right)^3$, $\left(x'_{N,i} \hat{\beta}_N\right)^2$ и $\left(x'_{N,i} \hat{\beta}_N\right)^3$. Они добав-

лялись в исходную модель в качестве объясняющих переменных: $\left(x'_{E,i} \hat{\beta}_E\right)^2$ и

$\left(x'_{E,i} \hat{\beta}_E\right)^3$ — в уравнение выхода в занятость, $\left(x'_{N,i} \hat{\beta}_N\right)^2$ и $\left(x'_{N,i} \hat{\beta}_N\right)^3$ — в уравнение выхода из рабочей силы. После этого проверялась их значимость (в совокупности) с помощью критерия отношения правдоподобия.

Для проверки была выбрана кусочно-постоянная модель как обеспечивающая лучшее качество подгонки. Статистика отношения правдоподобия оказалась равна 4,2186, а соответствующее p -значение – 0,38. Таким образом, проведенный тест не дает свидетельств против спецификации модели.

Кроме того, мы учитывали, что включение в модель индивидов, для которых время увольнения с последней работы неизвестно, может привести к смещению оценок. Для таких безработных мы не можем учесть временную зависимость, но, вероятно, она существует. В то же время, убрав такие наблюдения, мы теряем значительную часть выборки (393 наблюдения) и не можем оценить влияние опыта работы на продолжительность безработицы.

В Приложении 2 приводятся результаты оценки модели длительности безработицы как для всей анализируемой подвыборки, так и для ее укороченного варианта, где исключены наблюдения с неустанавливаемым сроком начала безработицы. Оценки как коэффициентов, так и временной зависимости оказываются близкими, что свидетельствует в пользу полученных результатов.

Заключение

В работе предложена эконометрическая модель продолжительности безработицы, которая может быть оценена на данных последовательности пространственных выборок (таких как данные РМЭЗ), содержащих повторяющиеся наблюдения за безработными индивидами. Модель учитывает различные варианты выхода из состояния безработицы, ее параметры интерпретируемы.

Проведено оценивание модели по данным РМЭЗ за 2000–2005 гг. Полученные оценки являются устойчивыми к изменениям спецификации, а также к изменению состава выборки (о чем свидетельствует близость оценок модели с учетом и без учета индивидов с неизвестным сроком начала безработицы).

Выявлена отрицательная временная зависимость для вероятности нахождения работы – по мере роста продолжительности пребывания в безработице вероятность трудоустройства падает. Этот вывод противоречит результатам Фолея [8], Гроган и ван ден Берга [9], на которых, видимо, сказалось смещение отбора (см. [6]). Оценки для вероятности выхода из состава рабочей силы свидетельствуют о немонотонной временной зависимости – эта вероятность сначала падает, а затем растет. Отметим, однако, что для получения более надежных выводов о характере временной зависимости стоит учитывать ненаблюдаемую разнородность индивидов, их «индивидуальный эффект». Тем не менее учет ненаблюдаемой разнородности сопряжен со значительными трудностями (проблемой начальных условий), и в данной работе индивидуальный эффект учитывался только простейшей моделью с разнородностью типа «mover-stayer» и отсутствием временной зависимости.

* *

*

СПИСОК ЛИТЕРАТУРЫ

1. Длительное пребывание в регистре безработных: низкий уровень образования, неудачное стечение обстоятельств или что-то еще? / Денисова И.А., Донецкий А.М., Колесникова О.А. и др. // Социальная политика: реалии XXI века. 2003. Вып. 1. С. 73–102.

2. Карцева М.А. Влияние образования на вероятность нахождения работы в России: Препринт РЭШ. BSP/2002/058R.
3. Ниворожжина Л.И., Ниворожкин Е.М., Шухмин А.Г. Моделирование поведения населения на рынке труда крупного города: продолжительность регистрируемой безработицы: Финальный отчет по гранту РПЭИ. 2000.
4. Ратникова Т.А. Введение в эконометрический анализ панельных данных // Экономический журнал ВШЭ. 2005. Т. 10. № 2. С. 267–316.
5. Ратникова Т.А., Фурманов К.К. Сколь долгов поиск работы в России: новый подход к оцениванию // Экономический журнал ВШЭ. 2009. Т. 13. № 2. С. 279–294.
6. Фурманов К.К. Об измерении средней продолжительности безработицы по данным Российского мониторинга экономики и здоровья // Прикладная эконометрика. 2009. № 2 (14). С. 74–99.
7. Brüderl J. Applied Regression Analysis Using Stata. 2005. <http://www.sowi.uni-mannheim.de/lehrstuehle/lessm/veranst/AppliedRegression.pdf>
8. Foley M.C. Determinants of Unemployment Duration in Russia. Yale University, 1997.
9. Grogan L., Berg G.J. van den. The Duration of Unemployment in Russia. University of Amsterdam, 1999.
10. Heckman J.J. The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete-Time Discrete-Data Stochastic Process // Structural Analysis of Discrete Data with Econometric Applications. Cambridge: MIT Press, 1981. P. 179–195.
11. Hilbe J.M. Logistic Regression Models. CRC Press, 2009.
12. Lancaster T. Econometric Methods for the Duration of Unemployment / *Econometrica*. 1979. Vol. 47. № 4. P. 939–956.
13. Nickell S.J. Estimating the Probability of Leaving Unemployment // *Econometrica*. 1979. Vol. 47. № 5. P. 1249–1266.
14. Schmidt P., Witte A.D. Predicting Criminal Recidivism Using «Split Population» Survival Time Models // *Journal of Econometrics*. 1989. Vol. 40. № 1. P. 141–159.

Приложение 1.

**Оценки коэффициентов модели длительности безработицы
для различных спецификаций функций временной зависимости Φ_E и Φ_N .**

	Без временной зависимости		Mover-Stayer		Квадратичная		Кусочно-постоянная		Кусочно-линейная	
	β	e^β	β	e^β	β	e^β	β	e^β	β	e^β
Уравнение выхода в занятость										
Половозрастные характеристики										
Пол (1 – женщина)	-0,12 (0,11)	0,89	-0,09 (0,12)	0,91	-0,11 (0,11)	0,90	-0,12 (0,11)	0,89	-0,12 (0,11)	0,89
Возраст >=50 лет	-0,41*** (0,13)	0,67	-0,49*** (0,14)	0,61	-0,40*** (0,13)	0,67	-0,40*** (0,13)	0,67	-0,38*** (0,13)	0,68
Возраст <=24 лет	0,66*** (0,10)	1,94	0,61*** (0,12)	1,83	0,53*** (0,10)	1,70	0,55*** (0,10)	1,73	0,55*** (0,11)	1,74
Образование										
Высшее	0,12 (0,12)	1,13	0,09 (0,13)	1,09	0,07 (0,12)	1,07	0,08 (0,12)	1,09	0,08 (0,12)	1,08
Ниже среднего	-0,01 (0,10)	0,99	-0,03 (0,10)	0,97	0,00 (0,10)	1,00	-0,01 (0,10)	0,99	-0,01 (0,10)	0,99
Положение на рынке труда										
Отсутствие опыта работы	-0,34 (0,25)	0,71	-0,32 (0,26)	0,73	-0,23 (0,25)	0,79	-0,25 (0,25)	0,78	-0,26 (0,25)	0,77
Отсутствие опыта рабо- ты × Пол	0,53** (0,25)	1,69	0,52** (0,27)	1,67	0,51** (0,25)	1,67	0,52** (0,25)	1,68	0,52** (0,25)	1,68
Регистрация в ГСЗ	0,13 (0,11)	1,14	0,09 (0,12)	1,10	0,06 (0,12)	1,06	0,08 (0,12)	1,09	0,08 (0,12)	1,08
Семейное положение										
Состоит в браке	0,06 (0,11)	1,07	0,16 (0,13)	1,17	0,10 (0,12)	1,11	0,08 (0,12)	1,09	0,08 (0,12)	1,09
Брак × Пол	-0,21 (0,15)	0,81	-0,33* (0,17)	0,72	-0,17 (0,16)	0,85	-0,16 (0,16)	0,85	-0,16 (0,16)	0,85
Субъективная оценка богатства										
Бедный	-0,42*** (0,09)	0,66	-0,43*** (0,10)	0,65	-0,38*** (0,09)	0,68	-0,39*** (0,09)	0,68	-0,38*** (0,09)	0,68
Богатый	-0,04 (0,14)	0,96	-0,03 (0,15)	0,97	-0,01 (0,14)	0,99	-0,01 (0,14)	0,99	0,00 (0,14)	1,00

Продолжение табл.

	Без временной зависимости		Mover-Stayer		Квадратичная		Кусочно-постоянная		Кусочно-линейная	
	β	e^β	β	e^β	β	e^β	β	e^β	β	e^β
Географические характеристики (базовая категория – областной центр)										
Город	-0,18* (0,10)	0,84	-0,21* (0,11)	0,81	-0,13 (0,10)	0,88	-0,12 (0,10)	0,89	-0,13 (0,10)	0,88
ПГТ	-1,11*** (0,17)	0,33	-1,21*** (0,19)	0,30	-1,03*** (0,18)	0,36	-1,01*** (0,18)	0,36	-1,01*** (0,18)	0,36
Село	-0,36*** (0,10)	0,70	-0,42*** (0,12)	0,66	-0,34*** (0,11)	0,71	-0,34*** (0,11)	0,71	-0,35*** (0,11)	0,71
Москва – Санкт-Петербург	-0,52*** (0,16)	0,60	-0,59*** (0,17)	0,55	-0,50*** (0,16)	0,61	-0,52*** (0,16)	0,60	-0,50*** (0,16)	0,61
Региональный уровень безработицы	-0,05*** (0,01)	0,95	-0,05*** (0,01)	0,95	-0,04*** (0,01)	0,96	-0,04*** (0,01)	0,96	-0,04*** (0,01)	0,96
Дополнительно										
Не указано время увольнения	-0,59*** (0,16)	0,56	-0,60*** (0,17)	0,55	-0,99*** (0,17)	0,37	-1,75*** (0,45)	0,17	-1,86** (0,80)	0,16
Константа	-1,45 (0,15)	0,23	-1,24 (0,17)	0,29	-1,13 (0,16)	0,32	-0,36 (0,45)	0,70	-0,16 (0,80)	0,86
Уравнение выхода из состава рабочей силы										
Половозрастные характеристики										
Пол (1 – женщина)	-0,05 (0,14)	0,95	-0,04 (0,14)	0,96	-0,08 (0,14)	0,92	-0,08 (0,14)	0,92	-0,07 (0,14)	0,93
Возраст >=50 лет	0,47*** (0,13)	1,61	0,46*** (0,13)	1,58	0,45*** (0,13)	1,57	0,45*** (0,13)	1,57	0,45*** (0,13)	1,57
Возраст <=24 лет	-0,18 (0,13)	0,83	-0,18 (0,13)	0,83	-0,07 (0,14)	0,93	-0,07 (0,14)	0,93	-0,07 (0,14)	0,93
Образование										
Высшее	-0,07 (0,15)	0,93	-0,08 (0,15)	0,93	-0,03 (0,15)	0,97	-0,03 (0,15)	0,97	-0,03 (0,15)	0,97
Ниже среднего	-0,17 (0,12)	0,84	-0,18 (0,12)	0,84	-0,20 (0,12)	0,82	-0,20* (0,12)	0,82	-0,20 (0,12)	0,82
Положение на рынке труда										
Отсутствие опыта работы	-0,07 (0,25)	0,93	-0,07 (0,25)	0,94	-0,17 (0,25)	0,84	-0,16 (0,25)	0,85	-0,16 (0,25)	0,86
Отсутствие опыта работы × Пол	0,52** (0,25)	1,68	0,52** (0,26)	1,68	0,57** (0,26)	1,76	0,56** (0,26)	1,75	0,56** (0,26)	1,74
Регистрация в ГСЗ	-0,48*** (0,16)	0,62	-0,48*** (0,16)	0,62	-0,42*** (0,16)	0,66	-0,40** (0,16)	0,67	-0,42*** (0,16)	0,66

Окончание табл.

	Без временной зависимости		Mover-Stayer		Квадратичная		Кусочно-постоянная		Кусочно-линейная	
	β	e^β	β	e^β	β	e^β	β	e^β	β	e^β
Семейное положение										
Состоит в браке	-0,20 (0,14)	0,82	-0,19 (0,14)	0,83	-0,24* (0,14)	0,78	-0,23 (0,14)	0,80	-0,23 (0,14)	0,79
Брак × пол	0,36* (0,18)	1,43	0,34* (0,19)	1,40	0,35* (0,19)	1,42	0,35* (0,19)	1,41	0,35* (0,19)	1,42
Субъективная оценка богатства										
Бедный	-0,01 (0,10)	0,99	-0,02 (0,10)	0,99	-0,03 (0,11)	0,97	-0,03 (0,11)	0,97	-0,03 (0,11)	0,97
Богатый	0,18 (0,15)	1,20	0,18 (0,15)	1,20	0,17 (0,15)	1,19	0,17 (0,15)	1,18	0,18 (0,16)	1,19
Географические характеристики (базовая категория – областной центр)										
Город	0,00 (0,13)	1,00	-0,02 (0,14)	0,98	-0,06 (0,14)	0,94	-0,07 (0,14)	0,94	-0,06 (0,14)	0,94
ПГТ	0,31* (0,16)	1,37	0,29* (0,17)	1,34	0,24 (0,17)	1,27	0,23 (0,17)	1,25	0,24 (0,17)	1,26
Село	0,47*** (0,13)	1,60	0,45*** (0,13)	1,57	0,42*** (0,13)	1,52	0,42*** (0,13)	1,51	0,43*** (0,13)	1,53
Москва – Санкт-Петербург	0,29 (0,20)	1,33	0,26 (0,20)	1,30	0,27 (0,20)	1,31	0,28 (0,20)	1,32	0,29 (0,20)	1,33
Региональный уровень безработицы	0,02* (0,01)	1,02	0,02* (0,01)	1,02	0,02* (0,01)	1,02	0,02* (0,01)	1,02	0,02* (0,01)	1,02
Временной эффект										
2003 г.	0,25** (0,11)	1,29	0,25** (0,11)	1,28	0,24** (0,11)	1,27	0,25** (0,11)	1,29	0,25** (0,11)	1,28
2005 г.	0,34*** (0,11)	1,40	0,34*** (0,11)	1,40	0,34*** (0,12)	1,40	0,35*** (0,12)	1,42	0,35*** (0,12)	1,42
Дополнительно										
Не указано время увольнения	0,09 (0,16)	1,09	0,08 (0,16)	1,08	0,44** (0,18)	1,55	-0,52 (0,44)	0,59	-0,56 (0,77)	0,57
Константа	-3,34 (0,18)	0,04	-3,30 (0,19)	0,04	-3,61 (0,20)	0,03	-2,64 (0,45)	0,07	-2,52 (0,78)	0,08
ln(L)	-1826,36		-1819,72		-1771,59		-1765,92		-1767,13	
AIC	3732,72		3721,44		3631,17		3627,84		3630,26	
N	2010		2010		2010		2010		2010	

Примечания. В скобках указаны оценки стандартных отклонений оценок коэффициентов.

* – значимость на уровне 10%, ** – на уровне 5%, *** – на уровне 1%.

Приложение 2.

**Оценки параметров модели длительности безработицы по полной
и укороченной выборке (из укороченной выборки исключались
респонденты, не имеющие опыта работы и не указавшие срок увольнения
с последнего места работы)**

Переменные	Выход в занятость		Выход из рабочей силы	
	полная выборка	короткая выборка	полная выборка	короткая выборка
Пол (1 – женщина)	-0,12 (0,11)	-0,10 (0,12)	-0,08 (0,14)	-0,07 (0,16)
Возраст:				
от 50 лет	-0,40 (0,13)	-0,45 (0,13)	0,45 (0,13)	0,48 (0,14)
до 24 лет	0,55 (0,10)	0,49 (0,11)	-0,07 (0,14)	0,01 (0,17)
Высшее образование	0,08 (0,12)	0,09 (0,13)	-0,03 (0,15)	-0,23 (0,19)
Образование ниже среднего	-0,01 (0,10)	0,05 (0,11)	-0,20 (0,12)	-0,22 (0,14)
Отсутствие опыта работы	-0,25 (0,25)		-0,16 (0,25)	
Отсутствие опыта работы × Пол	0,52 (0,25)		0,56 (0,26)	
Регистрация в ГСЗ	0,08 (0,12)	0,09 (0,13)	-0,40 (0,16)	-0,35 (0,18)
Состоит в браке	0,08 (0,12)	0,11 (0,12)	-0,23 (0,14)	-0,28 (0,16)
Брак × Пол	-0,16 (0,16)	-0,20 (0,17)	0,35 (0,19)	0,40 (0,21)
Бедный	-0,39 (0,09)	-0,33 (0,10)	-0,03 (0,11)	0,03 (0,12)
Богатый	-0,01 (0,14)	-0,08 (0,16)	0,17 (0,15)	0,08 (0,19)
Проживание:				
в городе	-0,12 (0,10)	-0,14 (0,11)	-0,07 (0,14)	-0,09 (0,16)
в ПГТ	-1,01 (0,18)	-0,97 (0,20)	0,23 (0,17)	0,21 (0,20)
в селе	-0,34 (0,11)	-0,37 (0,12)	0,42 (0,13)	0,42 (0,15)
Москва – Санкт-Петербург	-0,52 (0,16)	-0,40 (0,18)	0,28 (0,20)	0,49 (0,23)
Региональный уровень безработицы	-0,04 (0,01)	-0,02 (0,01)	0,02 (0,01)	0,04 (0,01)
Дамми для				
2003 г.			0,25 (0,11)	0,41 (0,13)
2005 г.			0,35 (0,12)	0,36 (0,13)
Не указано время увольнения	-1,75 (0,45)		-0,52 (0,44)	
Константа	-0,36 (0,45)	-0,53 (0,45)	-2,64 (0,45)	-2,91 (0,46)

Окончание табл.

Переменные	Выход в занятость		Выход из рабочей силы	
	полная выборка	короткая выборка	полная выборка	короткая выборка
Временная зависимость				
φ_1	-0,87 (0,51)	-0,87 (0,50)	-1,28 (0,59)	-1,30 (0,59)
φ_2	-1,26 (0,41)	-1,26 (0,40)	-0,64 (0,41)	-0,66 (0,41)
φ_3	-1,29 (0,45)	-1,30 (0,45)	-0,48 (0,46)	-0,50 (0,46)
φ_4	-1,69 (0,43)	-1,72 (0,43)	-0,28 (0,42)	-0,29 (0,42)
N	2010	1617	2010	1617