

УДК 331.2+ 330.43

*А.Ю. Мерзлякова***ФАКТОРЫ ВЕЛИЧИНЫ ИНДИВИДУАЛЬНОЙ ЗАРАБОТНОЙ ПЛАТЫ**

Используются панельные микроданные Российского мониторинга экономического положения и здоровья населения за 2006-2015 гг. Анализируется каким образом влияют на величину заработной платы индивида такие факторы, как гендерные различия, уровень образованности, стаж работы и отрасль, в которой занят индивид. На основе статистического анализа показано, что в среднем при прочих равных условиях заработная плата женщины меньше, чем у мужчин. Проверена гипотеза о влиянии «инвестиций в человеческий капитал» на уровень заработной платы. Предложенный анализ использует несколько методологических приемов. В первой части при выявлении гендерных отличий в уровне заработной платы, а также при анализе отраслевых различий используются методы регрессионного анализа. А при анализе влияния уровня образования на заработную плату во второй части статьи используются модели панельных данных, которые позволяют устранить эффект эндогенности и учесть особенности индивида, которые являются ненаблюдаемыми факторами. К тому же с помощью учета перекрестных взаимодействий удалось выявить экономическую отдачу от образования¹, которая тесно связана с временной компонентой, и в последние годы имеет тенденцию к росту.

Ключевые слова: факторы заработной платы, эконометрическое моделирование, модель панельных данных.

Основной экономической теорией, определяющей детерминанты заработной платы, является теория человеческого капитала. В ее развитие внесли важный вклад Дж. Минцер (Jacob Mincer, 1957–1962), Т. Шульц (Theodore Shultz, 1960–1961) и Г. Беккер (Gary Becker, 1962–1964). Современная теория человеческого капитала уходит своими корнями в классическое произведение Адама Смита об уравнивании различий в оплате труда, созданное им и XVIII в.

Г. Беккер [1] и Дж. Минцер [2] рассматривали особенности повышения квалификации как форму приобретения человеческого капитала с точки зрения заработной платы и предложения труда.

Если рассматривать работы Э. Лейзера (Edward Lazear, 1981) [3], то он предполагал, что фирмы могут создавать компенсационные схемы «зарплата-опыт» для работников со специальными навыками, обеспечивая им стимулы, совместимые с интересами фирмы.

Г. Гец и Г. Беккер (Gilbert Ghez and Gary Becker, 1974) [4] рассматривали как количество часов работы и доходы индивида изменяются в течение жизни. Получаемые результаты говорят о том, что с увеличением возраста уменьшается отдача от инвестиций и желательно оптимально увеличивать рабочие часы в раннем возрасте. Они показывают, что оптимальная структура доходов, выстроенная по возрасту, имеет пик рабочих часов перед достижением максимума ставки заработной платы и доходов. Однако Дж. Хекман (James Heckman, 1976) [5], М. Киллингсворт (Mark Killingsworth, 1982) [6] и Й. Вайс (Yoram Weiss, 1986) [7] подчеркивают, что оптимальная структура формы и пиков жизненных циклов может меняться при получении специфических навыков или при очень большой ставке дисконтирования.

В своих работах К. Эрроу [8] (Kenneth Arrow, 1973) и М. Спенс [9] (Michael Spence, 1973-1974) выдвигают гипотезу, что образование не увеличивает производительность труда, но используется нанимателями как фильтр для приема на работу. Таким образом, люди с более высоким уровнем образования имеют больше шансов получить высокооплачиваемую работу, а менее обученные специалисты отсеиваются с таких должностей даже при равных способностях. Они доказывают, что даже если все люди с дифференцированными способностями осознают, что наниматели используют диплом как фильтр, вероятнее всего, их обладателями будут те, кому они обойдутся с наименьшими личными издержками времени и средств.

Таким образом, рассмотрим модель индивидуальной заработной платы, учитывающей все рассмотренные выше нюансы.

Для построения модели индивидуальной заработной платы используем данные Российского мониторинга экономического положения и здоровья населения (РМЭЗ) [10], которое представляет собой единственное в России представительное панельное обследование семей.

¹ Отдача от образования – это экономический эффект, который показывает на сколько увеличится заработная плата при изменении уровня образования.

В панель, которую будем исследовать, включены данные за 2006–2015 г. с шагом 2 года. Таким образом, в выборку входят 31395 наблюдений, из которых 44,2 % – мужчин, и соответственно 55,8 % – женщин. Средняя заработная плата в этой выборке для мужчин составляет 18760,49 руб., а для женщин – 12921,85 руб. Предположим, что заработная плата является линейной комбинацией некоторой константы и переменных, в состав которых входит «пол респондента». Данная переменная является фиктивной, то есть принимает только два значения:

$$x_i = \begin{cases} 1, & \text{если респондент женщина,} \\ 0, & \text{если респондент мужчина.} \end{cases}$$

Данная переменная отражает отличия заработной платы у мужчин и женщин, однако они могут быть вызваны разными факторами, такими как: условия труда, уровень образования, разный стаж работы, различия в способностях индивидов.

Включим в нашу модель регрессии следующие переменные:

sex_i – пол респондента;

edu_i – время обучения (в годах);

$stag_i$ – стаж работы на данном рабочем месте (в годах).

Таким образом, модель описывает математическое ожидание месячной заработной платы респондента при заданных значениях пола, времени обучения и стажа.

$$wage_i = \beta_1 + \beta_2 \cdot sex_i + \beta_3 \cdot edu_i + \beta_4 \cdot stag_i + \varepsilon_i$$

В этом случае коэффициент β_2 при переменной sex интерпретируется как разность математических ожиданий заработной платы между мужчиной и женщиной с одним и тем же временем обучения и стажем работы. Все коэффициенты модели интерпретируются только при условии *ceteris paribus* (при прочих равных условиях), то есть коэффициент можно трактовать только при неизменности прочих факторов, включенных в модель. Оценивание модели методом наименьших квадратов приводит к следующим результатам (табл. 1).

Таблица 1

Результаты метода наименьших квадратов для уравнения индивидуальной заработной платы

Переменная	Коэффициент	Стандартная ошибка	Критерий Стьюдента	p-value
Константа	14306,08	187,88	76,14	$< 2 \cdot 10^{-16}$
sex	-6597,44	153,22	-43,06	$< 2 \cdot 10^{-16}$
edu	428,74	13,57	31,60	$< 2 \cdot 10^{-16}$
$stag$	-9,49	8,27	-1,15	0,251

Для женщин: $zar_plata_i = 7708.64 + 428.74 \cdot edu_i - 9.49 \cdot stag_i + \varepsilon_i$;

для мужчин: $zar_plata_i = 14306.08 + 428.74 \cdot edu_i - 9.49 \cdot stag_i + \varepsilon_i$.

Коэффициент при переменной sex теперь предполагает, что если сравнивать мужчин и женщин с одним и тем же уровнем образования и стажем работы, то разность математических ожиданий составит 6597,44 руб. по сравнению с 5838,6 руб. прежде. Данная разность, как и в предыдущей модели, является статистически высоко значимой ($t = -43.06$ при $p\text{-value} < 2 \cdot 10^{-16}$). Также мы не можем принять нулевую гипотезу о незначимости оценки коэффициента при переменной edu ($t = 31.6$ при $p\text{-value} < 2 \cdot 10^{-16}$). Таким образом, при условии одного и того же пола и стажа работы один год дополнительного обучения приводит к росту заработной платы на 428,74 руб. Совместная гипотеза о равенстве нулю всех трех коэффициентов регрессии также отклоняется, поскольку критерий Фишера принимает значение 833.3 ($p\text{-value} < 2 \cdot 10^{-16}$), при соответствующем 5 % критическом значении 2,60. Для данной модели множественный коэффициент детерминации составляет 0,0738, то есть построенная модель описывает 7,38 % выборочной вариации заработной платы против 4,43 % по предыдущей модели.

Анализ ошибки модели показывает, что есть некоторая квадратическая зависимость между расчетными значениями и ошибками модели. Это может быть связано с тем, что опыт работника влияет на его заработную плату нелинейно: после многих лет работы эффект дополнительного года оказывает отрицательное влияние на заработную плату. Чтобы проверить это предположение включим в нашу модель квадрат стажа работы. Результаты данной модели представлены в табл. 2.

Для женщин: $zar_plata_i = 7207.09 + 426.95 \cdot edu_i + 151.76 \cdot stag_i - 5.12 \cdot stag_i^2 + \varepsilon_i$;

для мужчин: $zar_plata_i = 13835.46 + 426.95 \cdot edu_i + 151.76 \cdot stag_i - 5.12 \cdot stag_i^2 + \varepsilon_i$.

Таблица 2

Результаты метода наименьших квадратов для уравнения индивидуальной заработной платы спецификации 2

Переменная	Коэффициент	Стандартная ошибка	Критерий Стьюдента	p-value
Константа	13835,46	197,05	70,21	$< 2 \cdot 10^{-16}$
<i>sex</i>	-6628,37	153,12	-43,29	$< 2 \cdot 10^{-16}$
<i>edu</i>	426,95	13,56	31,49	$< 2 \cdot 10^{-16}$
<i>stag</i>	151,76	22,15	6,85	$7.43 \cdot 10^{-12}$
<i>stag</i> ²	-5,12	0,65	-7,85	$4.38 \cdot 10^{-15}$
$R^2 = 0.0756$, $R^2 adj = 0.0755$, $F = 641.6$ ($p - value < 2 \cdot 10^{-16}$)				

Дополнительная переменная *stag*² как и ожидалось имеет отрицательное значение. При этом, в отличие от предыдущей спецификации модели, все оценки коэффициентов статистически значимы на любом доступном уровне. К тому же коэффициент детерминации увеличился до 7,56 %.

Однако и для данной модели ошибка не удовлетворяет условиям применения классического метода наименьших квадратов, в частности, нарушается условие постоянства дисперсии ошибки. Одним из способов устранить или уменьшить гетероскедстичность является изменение функционала модели. Например, вместо абсолютных значений заработной платы можно использовать её логарифмы.

Построим логлинейную модель, которая объясняет зависимость логарифма заработной платы от переменных полное число лет учебы и стаж работы. Результаты представлены в табл. 3.

Таблица 3

Оценки МНК для модели логарифма заработной платы

Переменная	Коэффициент	Стандартная ошибка	Критерий Стьюдента	p-value
Константа	9,2609	0,0193	480,5370	$< 2 \cdot 10^{-16}$
<i>sex</i>	-0,5240	0,0135	-38,7420	$< 2 \cdot 10^{-16}$
$\log(edu)$	0,1069	0,0068	15,7410	$< 2 \cdot 10^{-16}$
$\log(stag)$	0,0755	0,0183	4,1320	0,00004
$\log^2(stag)$	-0,0083	0,0057	-1,4620	0,144
$R^2 = 0.0503$, $R^2 adj = 0.0502$, $F = 415.3$ ($p - value < 2 \cdot 10^{-16}$)				

Для женщин:

$\log(zar_plata_i) = 8.7369 + 0.1069 \cdot \log(edu_i) + 0.0755 \cdot \log(stag_i) - 0.0083 \cdot \log^2(stag_i) + \varepsilon_i$;

для мужчин:

$\log(zar_plata_i) = 9.2609 + 0.1069 \cdot \log(edu_i) + 0.0755 \cdot \log(stag_i) - 0.0083 \cdot \log^2(stag_i) + \varepsilon_i$.

При построении логлинейной модели изменяется интерпретация коэффициентов модели, в частности коэффициент при времени обучения $\hat{\beta}_3 = 0.1069$ показывает, что при прочих равных условиях при увеличении времени обучения на 1 % ожидаемая заработная плата увеличивается приблизи-

тельно на 0,11 %. Разность между ожидаемой логарифмической заработной платой между женщинами и мужчинами равна $-0,5240$, это соответствует разности, равной 40,79 %, то есть женщины получают в среднем на 40,8 % меньше, чем мужчины.

До сих пор эффект² пола респондента предполагался постоянным, независимо от опыта или уровня образования.

Возможно, что мужчины вознаграждаются по-другому, нежели женщины, имеющие более высокое образование, в этом случае модели, рассмотренные ранее, могут быть ограниченными. Данную разницу в заработной плате можно учесть с помощью введения перекрестных дамми-переменных, которые являются произведением фиктивной переменной пола и другими объясняющими переменными, включенными в модель:

$$y_i = \beta_1 + \beta_2 \cdot d_i + \beta_3 \cdot u_{2i} + \beta_4 \cdot u_{3i} + \beta_5 \cdot u_{4i} + \beta_6 \cdot s_i + \beta_7 \cdot ud_{2i} + \beta_8 \cdot ud_{3i} + \beta_9 \cdot ud_{4i} + \beta_{10} \cdot sd_i + \varepsilon_i,$$

где y_i – логарифм заработной платы i – го респондента; s_i – стаж работы в годах;

$$d_i = \begin{cases} 1, & \text{если респондент мужчина,} \\ 0, & \text{если респондент женщина.} \end{cases};$$

$$u_{1i} = \begin{cases} 1, & \text{если образование ниже среднего,} \\ 0, & \text{иначе} \end{cases}$$

$$u_{2i} = \begin{cases} 1, & \text{если среднее образование,} \\ 0, & \text{иначе} \end{cases};$$

$$u_{3i} = \begin{cases} 1, & \text{если среднее – специальное образование,} \\ 0, & \text{иначе} \end{cases}$$

$$u_{4i} = \begin{cases} 1, & \text{если высшее образование,} \\ 0, & \text{иначе} \end{cases};$$

$$ud_{ji} = \begin{cases} u_{ji}, & \text{если респондент мужчина,} \\ 0, & \text{если респондент женщина.} \end{cases},$$

$$sd_i = \begin{cases} s_i, & \text{если } i\text{-ый респондент мужчина,} \\ 0, & \text{если } i\text{-ый респондент женщина.} \end{cases}$$

Оценка данной модели методом наименьших квадратов приводит к результатам, представленным в табл. 4.

Таблица 4

Оценки коэффициентов робастные к гетероскедестичности

Переменная		Коэффициент	Стандартная ошибка	Критерий Стьюдента	p-value
Константа		8,6231	0,0447	193,0250	0,0000
d	пол	0,6839	0,0500	13,6645	0,0000
u_2	СО	0,2068	0,0481	4,2994	0,0000
u_3	СПО	0,2994	0,0476	6,2921	0,0000
u_4	ВПО	0,6819	0,0479	14,2413	0,0000
s	Стаж	0,0058	0,0007	7,7967	0,0000
ud_2	Уровень образования с учетом пола	-0,1140	0,0544	-2,0955	0,0361
ud_3		-0,0457	0,0541	-0,8450	0,3981
ud_4		-0,1399	0,0547	-2,5549	0,0106
sd	Стаж с учетом пола	-0,0039	0,0011	-3,4532	0,0006
$R^2 = 0.0736, \quad R^2 adj = 0.0733, \quad F = 276.7$					

Для мужчин: $y_i = 9.3070 + 0.0928 \cdot u_{2i} + 0.2537 \cdot u_{3i} + 0.542 \cdot u_{4i} + 0.0019 \cdot s_i + \varepsilon_i$;

для женщин: $y_i = 8.6231 + 0.2068 \cdot u_{2i} + 0.2994 \cdot u_{3i} + 0.6819 \cdot u_{4i} + 0.0058 \cdot s_i + \varepsilon_i$.

Две из трех фиктивных переменных, отвечающие за образование во взаимодействии с дамми-переменной пола, значимы на 5 % уровне. Это свидетельствует об эффекте от образования, то есть у

² Эффект – это результат воздействия переменной на рассматриваемый фактор.

мужчин он ниже, чем у женщин. Аналогичный вывод можно сделать и для эффекта стажа, который для мужчин ниже, чем у женщин.

В данной модели разница между изменениями заработной платы у мужчин и женщин примерно 0,46 или 46 % при прочих равных условиях.

Рассмотрим теперь модель абсолютной заработной платы, в которую включаются переменные, отражающие отраслевую принадлежность респондента, относится ли производство к вредным или нет. Также есть предположение, что на уровень заработной платы оказывает влияние то, к какой форме собственности относится предприятие (с государственным участием или частное).

В результате в модель входят 18 фиктивных переменных по отрасли:

<i>otrasl1</i> –	Высокотехнологичные отрасли	<i>otrasl10</i> –	Органы управления
<i>otrasl2</i> –	Легкая, пищевая промышленность	<i>otrasl11</i> –	Образование
<i>otrasl3</i> –	Гражданское машиностроение	<i>otrasl12</i> –	Наука, культура
<i>otrasl4</i> –	Военно-промышленный комплекс	<i>otrasl13</i> –	Здравоохранение
<i>otrasl5</i> –	Нефтегазовая промышленность	<i>otrasl14</i> –	Армия, МВД, органы безопасности
<i>otrasl6</i> –	Другая отрасль тяжелой промышленности	<i>otrasl15</i> –	Торговля, бытовое обслуживание
<i>otrasl7</i> –	Строительство	<i>otrasl16</i> –	Финансы
<i>otrasl8</i> –	Транспорт, связь	<i>otrasl17</i> –	Энергетическая промышленность
<i>otrasl9</i> –	Сельское хозяйство	<i>otrasl18</i> –	Жилищно-коммунальное хозяйство

Соответственно, переменную, отражающую вредное производство или не отражающее, обозначим $vred_i$ ($1 = \text{"вредное"}; 0 = \text{"не вредное"}),$ компанию с государственным капиталом – gov_i ($1 = \text{"с государственным капиталом"}; 0 = \text{"иначе"}),$ а компанию с частным капиталом – $private_i$ ($1 = \text{"с частным капиталом"}; 0 = \text{"иначе"}).$

В результате получаем модель (табл. 5), которая описывает 18,68 % вариации индивидуальной заработной платы, что гораздо лучше всех выше описанных моделей.

Из табл. 5 видно, что при прочих равных условиях заработная плата мужчин больше заработной платы женщин на 4824,40 руб. Также можно увидеть, что отдача от среднего и средне-специального образования одинакова для мужчин и женщин (поскольку коэффициенты при данных переменных статистически не значимы, а значит можно считать, что они равны нулю), а вот отдача от высшего образования больше у мужчин, чем у женщин.

Так как коэффициенты при отраслевых переменных «Строительство», «Транспорт и связь», а также «Финансы» оказались статистически незначимы, то можно сделать вывод, что при прочих равных условиях уровень заработной платы в этих отраслях не отличается от среднего уровня, а лидерами в оплате труда являются такие отрасли, как нефтегазовая промышленность (больше среднего уровня на 6923,47 руб.) и высокотехнологичные отрасли (выше среднего уровня на 4902,30 руб.), а наименее оплачиваемые – образование (меньше среднего уровня на 5928,17 руб.), гражданское машиностроение (меньше на 6777,28 руб.) и сельское хозяйство (меньше на 9480,61 руб.).

Вывод: на предприятиях с государственным капиталом респонденты получают меньше, чем при участии частного капитала.

Таким образом, если респондент мужчина, имеющий высшее образование, работает в компании нефтяной промышленности (не вредное производство), которая в составе имеет частный капитал, а стаж работы составляет 10 лет, то прогноз его заработной платы составит:

$$y = 13022,77 + 4824,4 + (7948,19 + 2268,57) + 10 \times 70,05 + 6923,47 + 1423,74 = 37111,64$$

мужчина высшее образование стаж=10 нефтегаз пром частный капитал

При этом, если рассматривается среднестатистический мужчина³, то доверительный интервал прогноза равен [36174,79; 38048,49].

Если же мы имеем ввиду конкретного мужчину, то доверительный интервал прогноза примет вид [12706,81; 61516,48].

³ Среднестатистический мужчина, имеющий высшее образование и работающий в компании нефтяной промышленности, которая в составе имеет частный капитал – это мужчина в возрасте 35,5 лет, 90,8% из этой выборки проживает в городе, а 66,1% выборки состоит в браке.

Таблица 5

Модель с отраслевой структурой заработной платы

Переменная	Коэффициент	Стандартная ошибка	Критерий Стьюдента	p-value
константа	13022,77	1100,22	10,65	0,00
<i>d</i>	4824,40	336,41	14,34	0,00
<i>u</i> ₂	871,23	235,75	3,55	0,00
<i>u</i> ₃	2051,74	233,88	8,27	0,00
<i>u</i> ₄	7948,19	277,20	28,67	0,00
<i>s</i>	70,05	8,36	8,36	0,00
<i>ured</i>	1076,78	200,84	5,13	0,00
<i>otrasl1</i>	4902,30	1525,00	3,46	0,00
<i>otrasl2</i>	-4912,90	1075,51	-4,57	0,00
<i>otrasl3</i>	-6777,28	1098,06	-6,17	0,00
<i>otrasl4</i>	-4044,47	1133,29	-3,57	0,00
<i>otrasl5</i>	6923,47	1291,89	5,92	0,00
<i>otrasl6</i>	-4250,26	1102,85	-3,85	0,00
<i>otrasl7</i>	-1338,50	1090,63	-1,23	0,22
<i>otrasl8</i>	-1013,13	1075,61	-0,94	0,35
<i>otrasl9</i>	-9480,61	1083,43	-8,75	0,00
<i>otrasl10</i>	-3647,76	1113,49	-3,28	0,00
<i>otrasl11</i>	-5928,17	1064,58	-5,57	0,00
<i>otrasl12</i>	-4214,02	1114,09	-3,78	0,00
<i>otrasl13</i>	-3887,04	1075,36	-3,61	0,00
<i>otrasl14</i>	-2710,58	1094,60	-2,48	0,01
<i>otrasl15</i>	-3503,43	1068,64	-3,28	0,00
<i>otrasl16</i>	2026,01	1277,74	1,56	0,11
<i>otrasl17</i>	-3034,00	1137,57	-2,67	0,01
<i>otrasl18</i>	-5676,10	1079,08	-5,26	0,00
<i>gov</i>	-2943,53	293,57	-10,03	0,00
<i>private</i>	1423,74	298,99	4,18	0,00
<i>ud</i> ₂	121,43	368,58	0,33	0,74
<i>ud</i> ₃	594,50	406,02	1,42	0,14
<i>ud</i> ₄	2268,57	517,35	4,50	0,00
<i>sd</i>	-24,74	15,17	-1,63	0,10
$R^2 = 0.1868, \quad R^2_{adj} = 0.186, \quad F = 239,9 \quad p\text{-value} < 2.2 \cdot 10^{-16}$				

А теперь рассмотрим женщину с тем же уровнем образования и стажем, но работающую в сфере образования, тогда прогноз её заработной платы будет равен:

$$y = 13022,77 + \underset{\text{высшее образование}}{7948,19} + 10 \times \underset{\text{стаж}=10}{70,05} - 5928,17 - 2943,53 = 12799,76$$

Доверительный интервал прогноза для среднестатистической женщины⁴ будет равен [12334,56; 13264,97], а для конкретной женщины доверительный интервал прогноза составит [1591,52; 37191,05].

⁴ Среднестатистическая женщина с высшим образованием, работающая в сфере образования – это женщина в возрасте 42,8 лет, 69,6 % из этой выборки проживает в городе, а 57,9 % выборки состоит в браке.

На уровень заработной платы влияют также и личные факторы индивида, которые не поддаются прямой регистрации. Данные факторы можно учесть с помощью моделей панельных данных, которые учитывают как объектную, так и временную структуру данных.

Анализ панельных данных позволяет оценивать модели сквозных регрессий, а также модели с индивидуальными детерминированным или случайным эффектами.

В модели регрессии с детерминированным индивидуальным эффектом предполагается, что свободный член модели принимает различные значения для каждого объекта, чтобы отразить влияние пропущенных или ненаблюдаемых переменных, характеризующих индивидуальные особенности исследуемых объектов, не меняющиеся во времени.

Модель регрессии со случайным индивидуальным эффектом предполагает, что индивидуальные различия носят случайный характер и в среднем нивелируются, а их теоретические дисперсии предполагаются одинаковыми для всех объектов выборки. Эта модель является компромиссом между сквозной регрессией и моделью фиксированных эффектов, так как менее ограничена, чем первая модель, и позволяет получить более значимые оценки, чем вторая [11].

Исходными данными, как и в первой части этой главы, являются данные Российского мониторинга экономического положения и здоровья населения (РМЭЗ) [10] за период 2006–2014 г. с шагом 2 года.

Оцениваемое уравнение имеет вид:

$$\begin{aligned} zar_plata_i = & \beta_0 + \beta_1 \cdot edu_i + \beta_2 \cdot age_i + \beta_3 \cdot age_i^2 + \beta_4 \cdot stag_i + \beta_5 \cdot sex_i + \\ & + \beta_6 \cdot civil_st_i + \beta_7 \cdot city_i + \beta_8 \cdot prof_{2i} + \beta_9 \cdot prof_{3i} + \beta_{10} \cdot prof_{4i} , \\ & + \beta_{11} \cdot prof_{5i} + \beta_{12} \cdot prof_{6i} + \beta_{13} \cdot prof_{7i} + \beta_{14} \cdot prof_{8i} + \varepsilon_i \end{aligned}$$

где zar_plata_i – месячная заработная плата (руб.); edu_i – продолжительность обучения в годах; age_i – возраст (в годах); $stag_i$ – стаж работы на данном месте (годах); sex_i – пол респондента (1 = *мужчина*, 0 = *женщина*); $civil_st_i$ – семейное положение (1 = *в браке*, 0 = *иначе*); $city_i$ – тип места жительства (1 = *город*, 0 = *иначе*); $prof_{ji}$ ($j = 1, 2, \dots, 8$) – дамми – переменные для профессиональных групп, где в качестве референтной берется $prof_1$ = *неквалифицированные рабочие всех отраслей*.

Проведем сквозное оценивание уравнения нашей модели, не учитывающее панельную природу данных, что эквивалентно построению обычных МНК-оценок (табл. 6).

Значение коэффициента детерминации не велико (модель описывает около 16 % вариации результативного признака). Такое значение может быть связано с тем, что в модель не включается такой показатель, как отрасль, в которой занят респондент. В целом критерий Фишера показывает значимость уравнения, к тому же оценки всех переменных (за исключением семейного положения) оказываются значимыми даже на 1 % уровне.

Построим теперь регрессию со случайными индивидуальными эффектами. Она позволяет учесть изменчивость выборки в ковариационной матрице случайных ошибок (табл. 6). В данной модели интерпретация коэффициента детерминации не имеет смысла, поскольку модель оценивается программой с помощью обобщенного метода наименьших квадратов с учетом того, что ошибка вбирает в себя информацию о пропущенных переменных, к которым относится и такая ненаблюдаемая переменная, как способности респондента. Как и предыдущая модель, на уровне значимости 5 % модель является статистически значимой, хотя значимость коэффициента при переменной $stag$ упала по сравнению с предыдущей моделью, а оценка коэффициента при переменной $civil_st$ так и осталась незначима.

Оценивание этой модели проводится в предположении, что независимые переменные и ошибка модели, которая включает индивидуальный эффект, не коррелированы между собой. Данное предположение может не выполняться, поскольку индивидуальный эффект содержит такие показатели, как способности индивида, которые скорее всего коррелируют с переменными обучение и профессиональная группа. Другая отсутствующая переменная – отраслевая принадлежность респондента также может коррелировать с возрастом, обучением и профессиональной группой. Для того чтобы учесть указанные взаимосвязи, оценим модель с детерминированными индивидуальными эффектами (табл. 6).

Таблица 6

Модели панельных данных

Переменная	Сквозная регрессия		Модель случайных эффектов		Модель детерминированных эффектов	
	Коэффициент	Стандартная ошибка	Коэффициент	Стандартная ошибка	Коэффициент	Стандартная ошибка
Константа	-11023,00***	797,65	-19697***	955,57	—	—
<i>edu</i>	201,96***	14,54	209,93***	15,37	21,81	19,57
<i>age</i>	695,45***	37,33	1006,50***	45,48	2647,49***	77,56
<i>age</i> ²	-8,61***	0,44	-10,55***	0,53	-11,74***	0,91
<i>stag</i>	-31,29***	9,2	21,20***	10,34	11,09	13,32
<i>sex</i>	7109,70***	170,2	7664,70***	231,12	—	—
<i>civil_st</i>	-129,32	151,58	-162,84	174,56	170,51	240,2
<i>city</i>	4237,20***	163,05	4450,90***	240,79	-1022,56	3118,75
<i>proof</i> ₂	3861,60***	267,9	2844,80***	283,33	2134,41***	353,97
<i>prof</i> ₃	3910,00***	316,38	3853,20***	331,89	1764,89***	412,45
<i>prof</i> ₄	4430,90***	377,4	4020,70***	380,76	1741,64***	452,78
<i>prof</i> ₅	6698,70***	286,4	6134,40***	300,74	2336,15***	381,56
<i>prof</i> ₆	8102,70***	297,92	6913,40***	325,86	2722,44***	440,76
<i>prof</i> ₇	14662,00***	368,81	11953,00***	390,59	5650,05***	490,64
<i>prof</i> ₈	6554,40***	1050,9	4881,70***	1072,6	3389,41***	1299,46
<i>R</i> ²	0,1593		0,1461		0,3266	
<i>F</i> -критерий	424,12		381,38		586,79	

В модели с детерминированными эффектами показатель детерминации увеличивается по сравнению со сквозной регрессией в 2 раза (с 0,16 до 0,33). Таким образом, можно сделать вывод, что большая часть вариации приходится на индивидуальные эффекты. Коэффициент при переменной *sex* не оценивается, так как данная переменная не изменяется во времени. Слабо меняются во времени и такие переменные, как *edu*, *civil_st*, *city*, что приводит к незначимым результатам. Но в целом модель можно признать статистически значимой. Также в модели отсутствует константа. Это связано с тем, что она включает в себя индивидуальный эффект, который принимает разные значения для разных респондентов.

Сравнив модели с помощью статистических тестов, приходим к тому, что индивидуальные эффекты являются статистически значимыми и их необходимо учитывать при анализе факторов, влияющих на размер индивидуальной заработной платы.

В предыдущих моделях мы не учли фактор времени, однако данные содержат периоды сильных потрясений в экономике России, например, кризис 2008 г., который мог повлиять на изменения заработной платы индивидуумов, включенных в обследование. Предположим, что временные эффекты имеют детерминированный характер, поэтому для их учета введем фактор времени в модель индивидуальных эффектов. Предположим, что эффект времени по-разному влиял на эффект образования. Для учета этого факта введем в модель перекрестные дамми-переменные, которые будут учитывать разный коэффициент наклона при образовании в разные годы (табл. 7).

Как видно из модели, действительно временной фактор очень важен для оценивания отдачи от образования. Можно сделать вывод, что с каждым годом отдача от образования увеличивается.

В целом по работе можно сделать следующие выводы. При построении статистических моделей подтвердился тот факт, что мужчины получают за свой труд больше, чем женщины при прочих равных условиях примерно на 46 %. К тому же включение таких переменных, как стаж работы и время обучения показало, что эффекты образования и стажа у женщин больше, чем у мужчин, а учет отраслевой принадлежности респондента указывает на значимые различия в выплатах заработных плат индивидуумов, работающих в разных отраслях.

Таблица 7

Модель с детерминированными индивидуальными эффектами с учетом временного фактора и разного наклона при образовании

Переменная	Коэффициенты	Стандартная ошибка	Критерий Стьюдента	p-value
<i>edu</i>	-126,44	24,72	-5,11	0,00
<i>age</i>	1299,51	245,47	5,29	0,00
<i>age</i> ²	-12,06	0,99	-12,24	0,00
<i>stag</i>	2,83	15,84	0,18	0,86
<i>civil _st</i>	116,26	285,73	0,41	0,68
<i>city</i>	-842,98	1554,59	-0,54	0,59
<i>proof</i> ₂	2048,97	361,87	5,66	0,00
<i>prof</i> ₃	1694,38	360,28	4,70	0,00
<i>prof</i> ₄	1781,40	410,08	4,34	0,00
<i>prof</i> ₅	2293,16	380,10	6,03	0,00
<i>prof</i> ₆	2700,22	498,73	5,41	0,00
<i>prof</i> ₇	5501,16	668,00	8,24	0,00
<i>prof</i> ₈	3674,73	1891,36	1,94	0,05
2008	2817,25	521,89	5,40	0,00
2010	3644,38	977,21	3,73	0,00
2012	5758,18	1443,56	3,99	0,00
2014	8304,91	1897,71	4,38	0,00
<i>edu</i> × 2008	105,01	21,98	4,78	0,00
<i>edu</i> × 2010	162,01	23,72	6,83	0,00
<i>edu</i> × 2012	217,12	30,75	7,06	0,00
<i>edu</i> × 2014	301,22	32,96	9,14	0,00
$R^2 = 0.3376, \quad R^2_{adj} = 0.3356, \quad F = 381.56, \quad p - value < 2.22 \cdot 10^{-16}$				

Так же как и в предыдущей модели существенную роль имеет пол респондента, а также входят в состав собственников государство или частные компании. Как правило, если респондент работает на вредном производстве, то ему предполагается доплата. В последних моделях удалось выявить, что на уровень заработной платы индивида влияют не только учтённые переменные, но и переменные, которые являются ненаблюдаемыми, например, способности респондента. Их удастся учесть с помощью модели детерминированных эффектов, которая показывает, что основное влияние на уровень заработной платы оказывает профессиональная принадлежность работника, то есть чем выше квалификация, тем выше заработная плата. К тому же удалось выявить отдачу от образования, которая тесно связана с временной компонентой и в последние годы имеет тенденцию к росту.

СПИСОК ЛИТЕРАТУРЫ

1. Becker Gary S. (1964). Human Capital: A Theoretical and Empirical Analysis, in Special Reference to Education, New York: National Bureau of Economic Research.
2. Mincer Jacob (1962). Labor Force Participation of Married Women: A Study of Labor Supply, in H. Gregg Lewis, ed.. Aspects of Labor Economics, Princeton, N.J.: Princeton University Press. P. 63-97.
3. Lazear Edward P. (1981). Agency, Earnings Profiles, Productivity, and Hours Restrictions, American Economic Review, 71.4, September. P. 606-620.
4. Ghez Gilbert R. and Gary S. Becker (1974). The Allocation of Time and Goods over the Life Cycle, New York: Columbia University Press for the National Bureau of Economic Research.
5. Heckman James J. (1976). A Life Cycle Model of Earnings, Learning and Consumption, Journal of Political Economy, 84:4, Part 2, August, S11-S44.

6. Killingsworth, Mark (1982). Learning by Doing and Investment in Training: A Synthesis of Two Rival Models of the Life Cycle, *Review of Economic Studies*, 49. P. 263-271.
7. Weiss Yoram (1986). The Determinants of Life Cycle Earnings: A Survey, Chapter 11 in Orley Ashenfelter and Richard Layard, eds., *The Handbook of Labor Economics*, Vol. 1, Amsterdam: North Holland-Elsevier Science Publishers. P. 603-639.
8. Arrow Kenneth J. (1973). Higher Education as a Filter, *Journal of Public Economics*, 2:3, August. P. 193-216.
9. Spence Michael A. (1974). *Market Signalling: Informational Transfer in Hiring and Related Screening Processes*. Cambridge, Mass.: Harvard University Press.
10. Российский мониторинг экономического положения и здоровья населения НИУ-ВШЭ (RLMS-HSE), проводимый Национальным исследовательским университетом «Высшая школа экономики» и ООО «Демоскоп» при участии Центра народонаселения Университета Северной Каролины в Чапел Хилле и Института социологии РАН. URL: <http://www.cpc.unc.edu/projects/rlms> и <http://www.hse.ru/rlms>.
11. Ратникова Т.А., Фурманов К.К. Анализ панельных данных и данных о длительности состояний: учеб. пособие / Национальный исследовательский университет «Высшая школа экономики». М., 2014. 373 с.

Поступила в редакцию 18.11.17

A.Yu. Merzlyakova

FACTORS OF THE SIZE OF INDIVIDUAL WAGES

Using the panel microdata of the Russian Monitoring of the Economic Situation and Health of Population for 2006-2015, the work analyzes how factors such as gender differences, educational level, length of service and the sector in which the individual is engaged affect the amount of the individual's wages. Based on statistical analysis, it is shown that, on average, all other things being equal, women receive fewer men. The hypothesis on the influence of investments in human capital on the level of wages has been tested. The proposed analysis uses several methodological methods. In the first part, with the identification of gender differences in the level of wages, as well as in the analysis of industry differences, regression analysis methods are used. And when analyzing the influence of the level of education on a payroll in the second part of the article, panel data models are used that can eliminate the effect of endogeneity and take into account the characteristics of the individual that are unobservable factors. In addition, through the integration of cross-interactions, it was possible to identify the impact of education, which is closely related to the time component, and has tended to grow in recent years.

Keywords: salary factors, econometric modeling, model of panel data.

Мерзлякова Анастасия Юрьевна,
кандидат экономических наук, доцент кафедры
математических методов в экономике
Института экономики и управления
ФГБОУ ВО «Удмуртский государственный университет»
426034, Россия, г. Ижевск, ул. Университетская, 1 (корп. 4)
E-mail: anastasiya_vavil@mail.ru

Merzlyakova A.Yu.,
Candidate of Economics, Associate Professor
at Department of Mathematical Methods in Economics
Institute of Economics and Management
Udmurt State University
Universitetskaya st., 1/4, Izhevsk, Russia, 426034
E-mail: anastasiya_vavil@mail.ru