

## **ФИНАНСОВЫЕ РИСКИ**

**Синявская Т.Г.,**

кандидат экономических наук, доцент, доцент кафедры «Математическая статистика, эконометрика и актуарные расчеты» Ростовского государственного экономического университета (РИНХ)

E-mail: sin-ta@yandex.ru

**Трегубова А.А.,**

кандидат экономических наук, доцент кафедры «Математическая статистика, эконометрика и актуарные расчеты» Ростовского государственного экономического университета (РИНХ)

E-mail: alexandra\_a\_t@mail.ru

### **ОЦЕНКА ВЗАИМОСВЯЗАННЫХ РИСКОВ В СТРАХОВАНИИ НА ОСНОВЕ МНОГОМЕРНЫХ ПРОБИТ-МОДЕЛЕЙ<sup>1</sup>**

Величина принимаемого страховой компанией риска зависит от числа наступивших страховых случаев (событий). В медицинском страховании страховой случай представляет собой обращение за медицинской помощью при возникновении у застрахованного индивида заболеваний, включенных в страховое покрытие. При этом заболевания могут наступать как по отдельности, так и одновременно в различных сочетаниях. Классические актуарные подходы, базирующиеся на расчете долей заболевших каждым типом болезни, не позволяют учесть возможных зависимостей заболеваний, когда одно является следствием другого или иным образом повышает вероятность его возникновения. Для корректной оценки страхового риска в этом случае могут быть использованы многомерные пробит-модели (multivariate probit). В статье приведены подходы к использованию моделей такого типа в медицинском страховании, оценены модели для нескольких заболеваний и выявлен ряд детерминант заболеваемости. Используются данные 22 волны Рос-

сийского мониторинга экономического положения и здоровья населения НИУ-ВШЭ (RLMS-HSE).

**Ключевые слова:** взаимосвязанные риски, корректировка страхового тарифа, медицинское страхование, многомерная пробит-модель, оценка рисков

**Sinyavskaya T.G., Tregubova A.A.**

### **DEPENDENT RISKS ASSESSMENT IN INSURANCE BASED ON MULTIVARIATE PROBIT MODELS**

The value of risk that has been taken on by an insurance company in health insurance depends on number of claims (number of insurance events) in the portfolio. Insured event in health insurance defined as an event, when insured due to illness has been provided with a health services. Herewith insured's diseases can occur simultaneously or in various combinations. Traditional actuarial methods based on calculation of proportion of insured with each disease among insured individuals don't take into account possible dependencies between diseases occurrence, when one is the result of another disease, or increases the probability of its occurrence. Multivariate probit models could provide an adequate evaluation of insurance risk in this case. This paper presents approaches to the multivariate probit models application in health insurance, model estimation results for a set of diseases and the main diseases determinants. Individual RLMS-HSE Round 22 data is used for model estimation.

**Keywords:** dependent risks, health insurance, insurance tariffs adjusting, multivariate probit, risk evaluation

#### **Введение**

Медицинское страхование представляет собой один из наиболее социально ориентированных видов. Целью медицинского страхования является предоставление населению гарантий получения медицинской помощи за счет накопленных страховых взносов, а также финансирования профилактических мероприятий. Однако в настоящее время в России обяза-

<sup>1</sup> Публикация подготовлена в рамках поддержанного РГНФ научного проекта № 16-02-00411

тельное медицинское страхование обеспечивает только минимальные гарантии из возможных. Более широкие возможности по получению медицинской помощи предоставляет добровольное медицинское страхование (ДМС).

Современная российская система добровольного медицинского страхования характеризуется значительной асимметрией спроса и предложения. Страховые компании заинтересованы в клиентах, не имеющих проблем со здоровьем в настоящее время, и зачастую выстраивают тарифные системы, заградительные для населения с высоким риском заболеваемости. В то время как страховать по ДМС в основном хотят индивиды, имеющие некоторые проблемы со здоровьем и желающие покрыть стоимость будущего лечения за счет полиса ДМС. В связи с этим убыточность данного вида страхования в целом по рынку остается на критическом уровне. Коэффициент выплат за период 2010-2014 гг. был довольно высоким, принимая значения 75-78%. При этом в 2014 году он составил 76,74%, что сопоставимо с уровнем 2010 года (76,59%). За первые 9 месяцев 2015 года уровень выплат в ДМС составил 67,72% [10]. Для решения этой проблемы страховым компаниям необходимо диверсифицировать портфель полисов ДМС за счет расчета тарифов для групп застрахованных с разным уровнем риска. Это требует разработки и применения адекватных методик оценки страхового риска, дающих возможность учитывать влияние различных рисков факторов и взаимную зависимость рисков [5, 3, 7].

Существует ряд подходов к формированию тарифов, базирующихся на альтернативных вариантах оценки и дооценки страхового риска. Одна из них применяется в страховании на случай смертельно опасных заболеваний и называется моделью с несколькими состояниями. Она была разработана Э. Дэш и Д. Гримшоу (1990) [9; 16] и предполагает рассмотрение трех состояний, в которые может попасть застрахованный: «здоровые», «больные смертельно опасным заболеванием» и «умершие». Для расчета страхового тари-

фа оцениваются переходные вероятности между различными состояниями [9, С.38]. Достоинством данной модели является высокая степень адекватности математического инструментария, что повышает точность получаемых оценок. К недостаткам относится сложность модели и недостаточность статистической базы [2, С.164], а также невозможность учета влияния факторов риска.

Иной подход, дающий возможность учесть неоднородность страхового риска в разных группах застрахованных, базируется на формировании системы поправочных коэффициентов. Одним из таких методов является система учета влияния категориальных рейтинговых факторов на величину страхового тарифа М.Ю. Жигалкина (2005) [1]. Под рейтинговыми факторами понимаются величины различной природы [1, С.40]. Это могут быть количественные переменные как дискретные, так и непрерывные (например, возраст застрахованного), и качественные как порядковые (группа здоровья застрахованного – «хорошее», «среднее», «плохое»), так и категориальные (например, пол застрахованного).

Другой метод получения поправочных коэффициентов основан на использовании моделей логистической регрессии (Синявская, Трегубова, 2011 [8]).

Тем не менее ни один из перечисленных методов не позволяет учесть потенциальных взаимосвязей между страхуемыми рисками. Например, в добровольном медицинском страховании риски состоят в возникновении у застрахованного одного или нескольких заболеваний, покрываемых страховым полисом. Вышеупомянутые методы дают возможность оценить влияние факторов как по отдельности, так и в комбинации, на возникновение каждого покрываемого заболевания, но не их сочетаний. На примере автострахования для решения такой задачи Valdez and Frees (2005) предложили, а Young, Valdez and Kohn (2006) доказали возможность использования многомерных пробит-моделей [18, 19]. В связи с тем что, как и в случае автострахования, в добровольном медицинском страховании отмечается

наличие взаимосвязанных рисков и реализация нескольких рисков при наступлении страхового случая, представляется актуальным изучить возможность применения многомерных пробит-моделей для оценки рисков ДМС.

### 1. Постановка задачи

Предположим, что имеется портфель страховых полисов, предусматривающий  $J$  возможных типов страховых претензий (заявлений о выплате страхового возмещения). Введем индикаторную переменную  $I_j$ , показывающую, что по договору из данного портфеля наступило страховое событие (предъявлена претензия) типа  $j$ , где  $j = 1, 2, \dots, J$ . Обозначим через  $M = (I_1, I_2, \dots, I_J)'$  вектор возможных типов страховых событий по данному полису. Можно представить, что в случае отсутствия страховых претензий данный вектор будет состоять из нулей. Таким образом, есть  $2^J$  возможных комбинаций векторов типов страховых событий  $M$ . Эта случайная переменная представляет собой дискретную случайную переменную с возможными значениями, описываемыми различными комбинациями  $I_j$ . Можно ограничить наши предположения теми наблюдениями, в которых возможны только  $2^j - 1$  типов страховых событий. Для целей исследования они названы компонентами условных типов страховых событий.

В актуарной литературе хорошо известно, что совокупное распределение страховых убытков получается его тради-

ционной декомпозицией на компоненты частот и тяжести убытков, где каждая компонента затем моделируется отдельно. Однако попыток моделирования совокупного распределения убытков для компонент условных типов страховых событий в актуарной литературе практически не предпринималось. Тем не менее недавно Valdez and Frees (2005) [18] предложили модель иерархической структуры для оценки и предсказания совокупного распределения страхового ущерба с использованием высокодетализированного набора данных на микроуровне. Для иллюстрации этой декомпозиции предположим, что есть класс риска  $i$  в календарном году  $t$ , в котором потенциально наблюдаемые результаты для единицы наблюдения  $\{it\}$  включают:

$K_{it}$  – индикатор страхового события в течение года;

$M_{it,j}$  – возможные типы страховых событий (комбинации претензий),  $j = 1, \dots, N_t$ ;

$A_{it,jk}$  – сумма (тяжесть) ущерба, возможная для каждого типа страхового события,  $j = 1, \dots, N_t$ , и для каждого типа страховых претензий  $k = 1, 2, \dots, m$ , где  $N_t$  – число страховых событий в календарном году  $t$ .

Объединенная функция распределения совокупного страхового ущерба может быть разложена как:

$$f(K, M, A) = f(K) \times f(M/K) \times f(A/K, M),$$

объединенная функция = частота  $\times$  условный тип страхового события  $\times$  условная тяжесть ущерба

где  $f(K, M, A)$  – объединенная функция распределения, равная результату компонент частоты, условного типа страховых событий и условной тяжести ущерба. В использованной нотации  $f(K)$  означает компоненту частоты, равную вероятности наличия (или отсутствия) страхового события в данном календарном году;  $f(M/K)$  – означает компоненту условного типа страхового события, равную вероят-

ности наступления события типа  $M$  для данного  $K$ ; и  $f(A/K, M)$  – означает компоненту условной тяжести ущерба, равную плотности вероятности для вектора  $A$ , для данных  $K$  и  $M$ . Здесь случайная переменная  $M$  описывает наблюдаемую комбинацию страховых претензий. Каждая наблюдаемая комбинация – это набор из  $m$  компонент в форме  $(i_1, i_2, \dots, i_m)$ , где каждое  $i_k$  для  $k = 1, 2, \dots, m$ , для каждого  $k$ -го типа

страховых претензий наблюдаемо или равно нулю в ином случае. Так, набор из  $m$  компонент  $(1, 0, \dots, 0)$  означает, что была предъявлена только страховая претензия, относящаяся к «первому типу страховых претензий», и аналогично,  $(1, 1, \dots, 1)$  означает предъявление страховых претензий всех типов.

## 2. Формулировка многомерной пробит-модели

Пусть  $I_j^0$  – ненаблюдаемый латентный отклик, ассоциируемый с  $j$ -м типом страхового события, для  $j = 1, 2, \dots, J$ , и  $I_j$  – означает бинарный отклик, соответствующий каждому типу страхового события. При использовании индикаторной функции  $I_j$  равняется единице, если наступило страховое событие типа  $j$ , и нулю, в противном случае. Таким образом, многомерная пробит-модель может быть специфицирована как линейная комбинация детерминированных и стохастических компонент следующим образом:

$$I_1^0 = x' \beta_1 + \epsilon_1, \text{ для } I_1 = I_{\{I_1^0 > 0\}}$$

$$I_2^0 = x' \beta_2 + \epsilon_2, \text{ для } I_2 = I_{\{I_2^0 > 0\}}$$

$$\vdots \quad \quad \quad \vdots$$

$$I_J^0 = x' \beta_J + \epsilon_J, \text{ для } I_J = I_{\{I_J^0 > 0\}}$$

где  $x = (1, x_1, \dots, x_p)'$  – вектор  $p$  ковариат, которые не различаются для разных типов страховых событий (детерминированная компонента), и  $\beta_j = (\beta_{j0}, \beta_{j1}, \dots, \beta_{jp})'$  – соответствующий вектор параметров, включая свободный член, подлежащий оценке. Отметим, что индекс наблюдения  $i$  был опущен для упрощения записи. Стохастическая компонента,  $\epsilon_j$ , может рассматриваться как состоящая из тех ненаблюдаемых факторов, которые объясняют предельную вероятность наступления страхового события типа  $j$ . Каждая  $\epsilon_j$  подчиняется  $J$ -мерному нормальному распределению с нулевым условным математическим ожиданием и дисперсией, нормированной единице для обеспечения идентифицируемости параметров, где  $\epsilon \sim N(0, \Sigma)$ , и ковариационная матрица  $\Sigma$  задана, как:

$$\Sigma = \begin{bmatrix} 1 & \rho_{12} \cdots & \rho_{1J} \\ \vdots & \ddots & \vdots \\ \rho_{J1} & \rho_{J2} \cdots & 1 \end{bmatrix}.$$

Особый интерес представляют элементы, не находящиеся на главной диагонали ковариационной матрицы,  $\rho_{sj}$ , которые представляют ненаблюдаемую корреляцию между стохастической компонентой  $s$ -го и  $j$ -го типов страховых событий. Более того, по причине симметричности ковариаций мы имеем  $\rho_{sj} = \rho_{js}$ .

Отметим, что в такой формулировке многомерной пробит-модели мы можем напрямую определить предельные эффекты. Например, предельная вероятность наблюдать  $j$ -й тип страхового события может быть выражена, как:

$$Pr(I_j = 1) = \Phi(x' \beta_j), \text{ для } j = 1, 2, \dots, J,$$

где  $\Phi(\cdot)$  – означает интегральную функцию стандартного нормального распределения. Более того, объединенная вероятность наблюдать все возможные типы страховых событий подчиняется  $J$ -мерному стандартному нормальному распределению:

$$Pr(I_1 = 1, \dots, I_J = 1) = \Phi_J(x' \beta_1, \dots, x' \beta_J; \Sigma),$$

где  $\Sigma$  – ковариационная матрица.

В оригинальной работе Ashford and Sowden (1970) [12] разработано понятие многомерной пробит-модели. Авторы обобщают одномерную пробит-модель бинарного отклика в предположении о многоуровневой, с векторными значениями, структуре отклика для различных физиологических систем организма. Квантовый отклик каждой системы рассматривается в качестве лежащей в основе непрерывной латентной переменной, которая дискретизируется для пороговой спецификации. После этой оригинальной работы многомерная пробит-модель была применена в различных контекстах в разных исследованиях, например, в работах Gibbons and Wilcox-Gök (1998), Balia and Jones (2004) [13, 17].

## 3. Описание данных для моделирования

Эмпирическая оценка возможности использования многомерных пробит-моделей для оценки рисков ДМС была

проведена на основе наличия у индивида заболеваний сердца, легких и позвоночника, рассмотренных как взаимозависимые и зависящих от ряда факторов. В связи с тем что адекватная поставленной задаче страховая статистика недоступна, для моделирования были использованы данные 22 волны Российского мониторинга экономического положения и здоровья населения НИУ-ВШЭ (RLMS-HSE) за 2013 год [11]. Объем сформированной выборки составил 10 102 индивида в возрасте от 18 лет до 69 лет включительно.

Выбор границ возраста аналогичен стандартным условиям медицинского страхования (обычно от 18 до 69 лет, реже – до 74 лет). Для целей моделирования количественная переменная возраста была разделена на группы (рис. 1), при этом последней группе соответствует укрупненный интервал 60-69 лет, так как принятие на страхование лиц старше 60 лет считается достаточно рискованным с точки зрения страховщика вне зависимости от точного возраста (как в возрасте 60 лет, так и в 65 лет, и так далее).

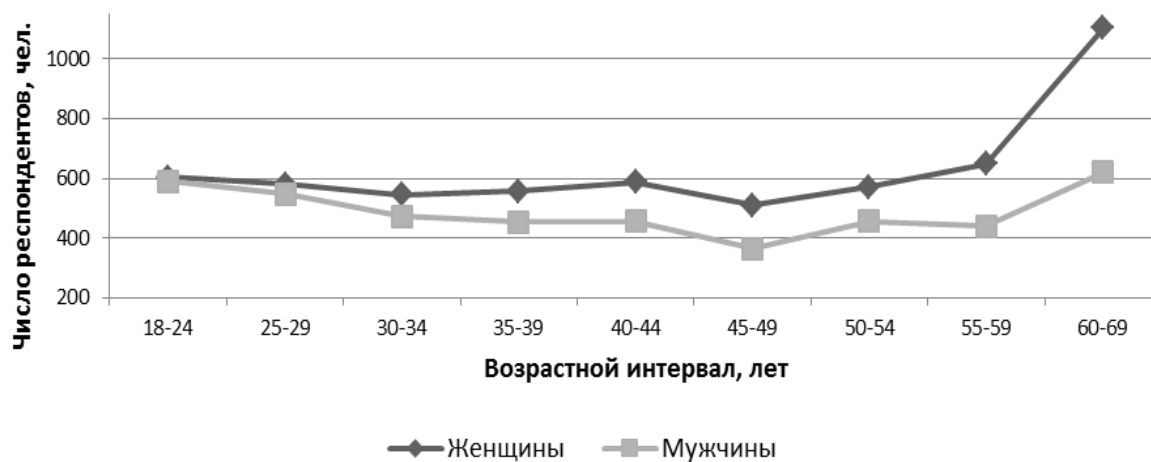


Рисунок 1 – Распределение респондентов по полу и возрасту, чел.

Доля мужчин в выборке составила 43,53%, женщин – 56,47%. Около 40% проживают в областных центрах, почти 27% – в сельской местности. Занимается трудовой деятельностью 65% индивидов, попавших в выборку (рис. 2). При этом более половины всех работающих мужчин (51,7%) являются квалифицированными рабочими, использующими машины и механизмы, а также занятыми ручным трудом. Неквалифицированных рабочих среди мужчин около 12,5%. Половина работающих женщин является специалистами высшего и среднего уровня квалификации, а также чиновниками, и почти 18% работает в сфере торговли и услуг.

Однако самооценка состояния здоровья женщинами и мужчинами различается

(рис. 3). Так, считают свое здоровье хорошим или даже очень хорошим почти половина мужчин (около 46,5%), и всего лишь 34,5% женщин. При этом оценивают свое здоровье, как среднее, чуть более 55% женщин и около 47% мужчин.

Отметим, что в младших возрастных группах представители обоих полов в основном оценивают свое здоровье как хорошее и очень хорошее, реже – как среднее (рис. 4). К 40-44 годам числа оценивающих свое здоровье как хорошее и как среднее становятся почти равными среди мужчин и женщин, а после 50 лет оценки закономерно меняются на средние и плохие, превращаясь после 60 лет, в основном, в негативные.

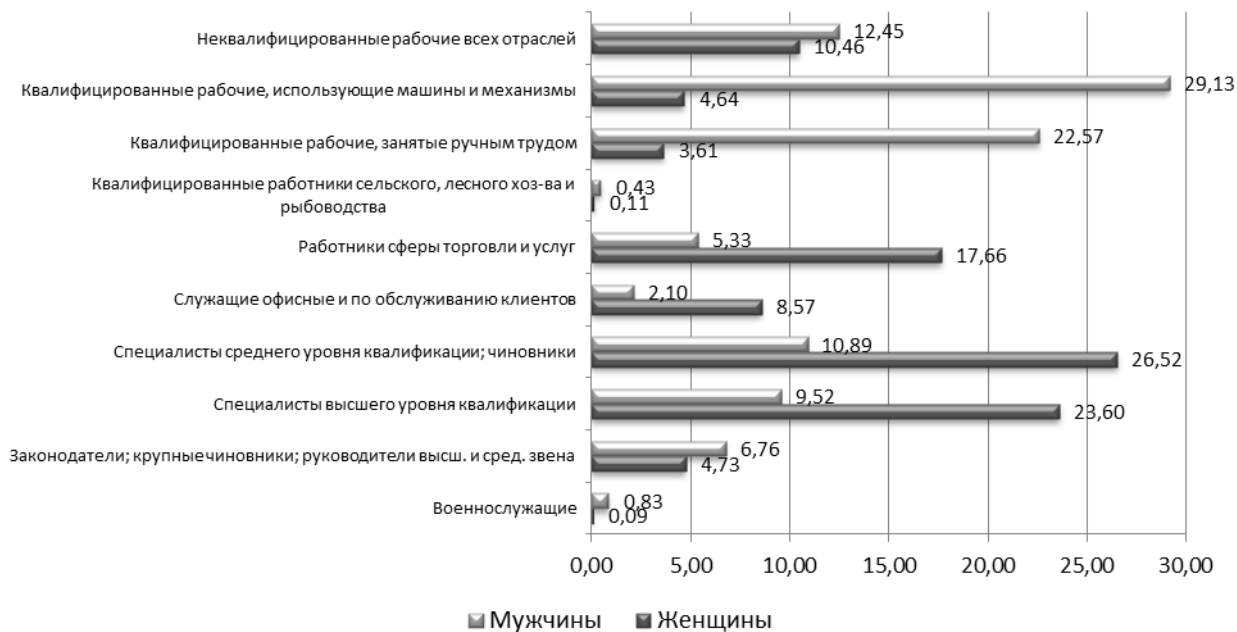


Рисунок 2 – Распределение респондентов по полу и профессиональной группе, %

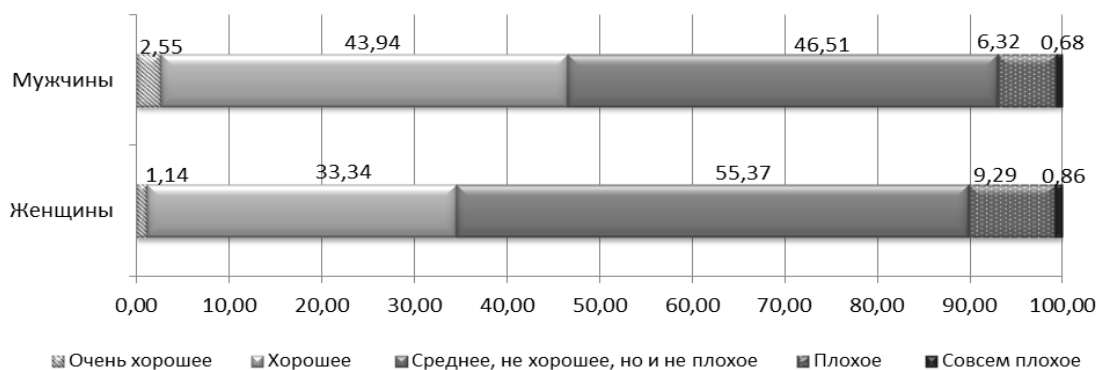


Рисунок 3 – Распределение оценок состояния здоровья по полу респондента, %

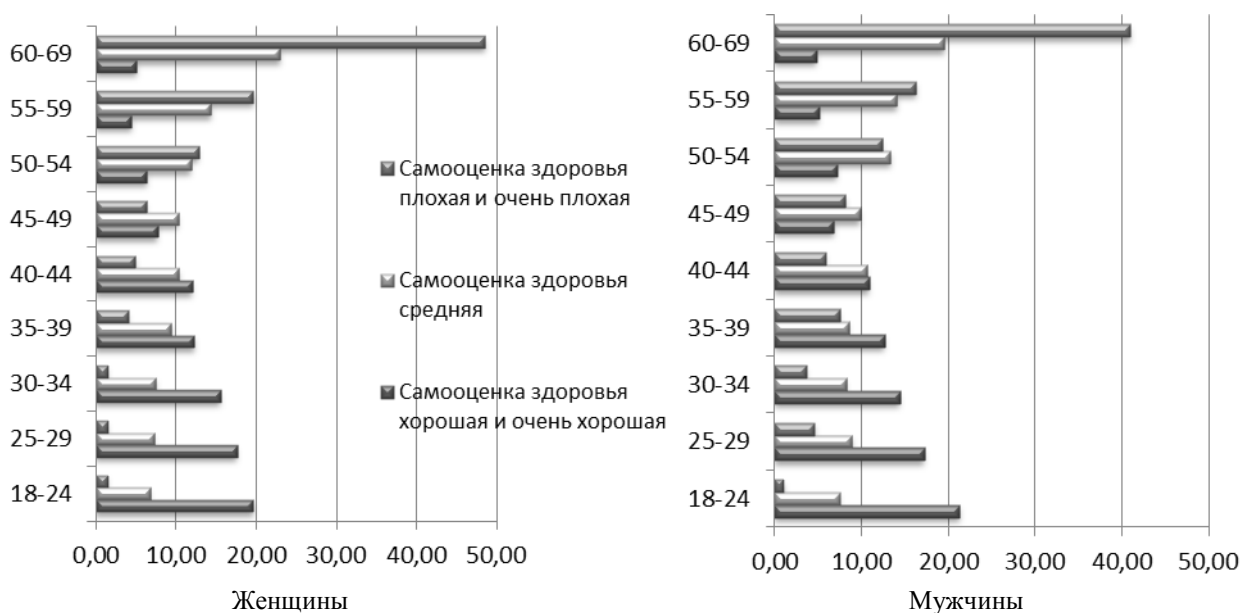


Рисунок 4 – Распределение оценок состояния здоровья по полу и возрасту респондента, %

#### 4. Статистический анализ взаимосвязанных заболеваний

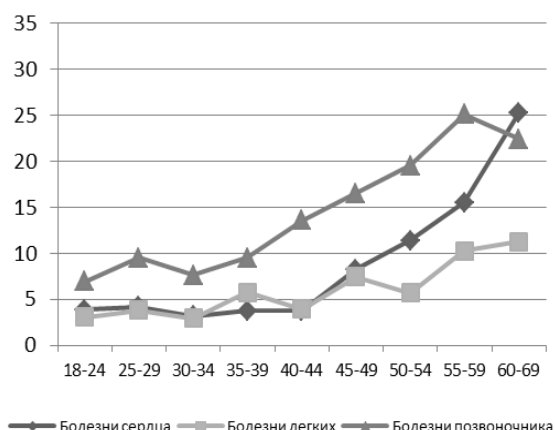
В связи с тем что цель данного исследования заключалась в анализе и моделировании взаимосвязанных рисков, был проведен анализ заболеваемости респондентов болезнями сердца, легких и позвоночника. Это обусловлено тем, что данные заболевания могут являться как сопутствующими (возникновение одного из этих заболеваний может сопровождаться наличием других), так вызывать друг друга (например, ишемическая болезнь сердца может приводить к хронической легочной недостаточности) [20]. Кроме того, данные заболевания могут иметь схожие симптомы (например, заболевания сердца и остеохондроз грудного отдела позвоночника сопровождаются болями в левой половине грудной клетки), что существенно осложняет диагностику и лечение, и даже может

привести к ошибкам в диагнозе, что для страховой компании может означать непредвиденный рост расходов [4, 6].

Доля лиц обоих полов, имеющих заболевания легких, колеблется в пределах 2-3% в младших возрастных группах и достигает 10-12% в старших (рис. 5). После 50 лет наблюдается значительный по сравнению с мужчинами рост доли страдающих заболеваниями сердца женщин (в группе 55-59 лет – 21% женщин против 15% мужчин, после 60 лет различие становится еще более заметным – 33,4% и 25,3% соответственно). Аналогично заболевания позвоночника в старших возрастных группах более распространены среди женщин по сравнению с мужчинами (в возрасте 55-59 лет, страдает 28% женщин против 25% мужчин, в возрасте 60-69 лет – 30,5% женщин против 22,4% мужчин).



Женщины



Мужчины

Рисунок 5 – Распределение заболеваний среди респондентов по полу и возрасту, в % к числу опрошенных в соответствующей возрастной группе

#### 5. Результаты моделирования

Детерминанты заболеваемости несколькими заболеваниями были выявлены на основе оценки параметров многомерной пробит-модель на основе метода имитационного максимального правдоподобия (англ. *simulated maximum likelihood, SML*) [14, 15] с тремя бинарными зависимыми переменными<sup>1</sup>. Для моделирования, проведенного по данным RLMS-HSE, были взяты следующие переменные:

(1) «У респондента есть заболевания сердца», принимающая значение, равное 1, в случае наличия заболевания и 0, при его отсутствии.

(2) «У респондента есть заболевания легких, бронхов», принимающая значение, равное 1, в случае наличия заболевания и 0, при его отсутствии.

(3) «У респондента есть заболевания позвоночника», принимающая значение, равное 1, в случае наличия заболевания и 0, при его отсутствии.

<sup>1</sup> Для расчетов использовался пакет STATA (11 версия), команда *mprobit* (Cappellari and Jenkins, 2003).

В качестве независимых выступили следующие характеристики индивидов: пол; уровень образования (сформированы четыре группы); статус занятости; наличие детей; тип населенного пункта; назначена ли инвалидность и получает ли индивид пенсию;

возрастная и доходная группы (квильные группы по доходу индивида), а также самооценка здоровья индивидов (субъективная оценка здоровья). Результаты оценки модели представлены в таблице 1.

Таблица 1 – Результаты оценивания многомерной пробит-модели

| Переменная   | Есть заболевания (1) сердца   |                    | Есть заболевания (2) легких |                    | Есть заболевания (3) позвоночника |                    |
|--|---|--------------------|-----------------------------|--------------------|-----------------------------------|--------------------|
|  | Коэффициент регрессии   | Стандартная ошибка | Коэффициент регрессии       | Стандартная ошибка | Коэффициент регрессии             | Стандартная ошибка |
| <b>Пол (женский)</b>   |   |                    |                             |                    |                                   |                    |
| Мужской  | -0,0533   | 0,0415             | 0,0101                      | 0,0443             | 0,0251                            | 0,0347             |
| <b>Тип поселения (областной центр)</b>                         |   |                    |                             |                    |                                   |                    |
| Город  | -0,0371   | 0,0493             | -0,0543                     | 0,0528             | -0,1823***                        | 0,0411             |
| ПГТ  | 0,3073***   | 0,0756             | -0,0072                     | 0,0883             | 0,0927                            | 0,0653             |
| Село   | 0,0977**  | 0,0495             | 0,0342                      | 0,0531             | -0,1024**                         | 0,0419             |
| <b>Уровень образования (высшее и послевузовское)</b>           |   |                    |                             |                    |                                   |                    |
| Неполное среднее (до 9 классов)                                | -0,1103   | 0,0697             | 0,0409                      | 0,0749             | -0,0645                           | 0,0590             |
| Среднее (9-11 классов)   | -0,0820   | 0,0533             | 0,0306                      | 0,0578             | -0,0645                           | 0,0441             |
| Среднее профессиональное                                       | -0,0433   | 0,0551             | 0,0413                      | 0,0601             | -0,0281                           | 0,0457             |
| <b>Статус занятости (неработающий)</b>                         |   |                    |                             |                    |                                   |                    |
| Работающий   | -0,1876***  | 0,0545             | 0,0790                      | 0,0595             | -0,0129                           | 0,0471             |
| <b>Наличие детей (нет)</b>                                     |   |                    |                             |                    |                                   |                    |
| Есть дети  | 0,0137  | 0,0634             | -0,1681***                  | 0,0628             | 0,0559                            | 0,0515             |
| <b>Получает пенсию (нет)</b>                                   |   |                    |                             |                    |                                   |                    |
| Получает пенсию  | 0,0310  | 0,0699             | -0,1339*                    | 0,0796             | 0,0625                            | 0,0606             |
| <b>Назначена инвалидность (нет)</b>                            |   |                    |                             |                    |                                   |                    |
| Назначена инвалидность   | 0,4831***   | 0,0660             | 0,2508***                   | 0,0773             | 0,0945                            | 0,0642             |
| <b>Возрастная группа (18-24 лет)</b>                           |   |                    |                             |                    |                                   |                    |
| 25-29 лет  | -0,0483   | 0,1131             | 0,1008                      | 0,1104             | -0,0472                           | 0,0874             |
| 30-34 лет  | -0,1509   | 0,1247             | 0,0637                      | 0,1203             | -0,0889                           | 0,0938             |
| 35-39 лет  | -0,1050   | 0,1202             | 0,2484**                    | 0,1147             | 0,0203                            | 0,0921             |
| 40-44 лет  | -0,0246   | 0,1172             | 0,1489                      | 0,1181             | 0,1542*                           | 0,0908             |
| 45-49 лет  | 0,2380**  | 0,1134             | 0,3610***                   | 0,1156             | 0,2601***                         | 0,0920             |
| 50-54 лет  | 0,2689**  | 0,1101             | 0,2891**                    | 0,1140             | 0,3577***                         | 0,0892             |
| 55-59 лет  | 0,4366***   | 0,1134             | 0,4784***                   | 0,1182             | 0,4419***                         | 0,0934             |
| 60-69 лет  | 0,6455***   | 0,1176             | 0,5545***                   | 0,1256             | 0,3629***                         | 0,0986             |
| <b>Самооценка состояния здоровья (плохое или очень плохое)</b> |   |                    |                             |                    |                                   |                    |
| Хорошее или очень хорошее                                      | -1,5128***  | 0,0749             | -1,0039***                  | 0,0787             | -1,0544***                        | 0,0622             |
| Среднее  | -0,7017***  | 0,0525             | -0,3789***                  | 0,0611             | -0,4053***                        | 0,0514             |
| <b>Квильти по доходу индивида (I)</b>                          |   |                    |                             |                    |                                   |                    |
| II   | -0,0559   | 0,0649             | -0,1863***                  | 0,0708             | 0,0565                            | 0,0569             |
| III  | -0,0871   | 0,0691             | -0,0831                     | 0,0724             | 0,0653                            | 0,0593             |
| IV   | -0,0343   | 0,0779             | -0,2045**                   | 0,0822             | 0,0455                            | 0,0650             |
| V  | -0,0654   | 0,0821             | -0,0583                     | 0,0827             | 0,1325**                          | 0,0667             |
| Константа  | -0,5813***  | 0,1075             | -1,1543***                  | 0,1151             | -0,6943***                        | 0,0917             |
| <b>Коэффициент корреляции терминов ошибок</b>                  |   |                    |                             |                    |                                   |                    |
| $\hat{\rho}_{21}$  | 0,0802***   |                    |                             |                    |                                   |                    |
| $\hat{\rho}_{31}$  | 0,1215***   |                    |                             |                    |                                   |                    |
| $\hat{\rho}_{32}$  | 0,0952***   |                    |                             |                    |                                   |                    |
| -2LL   | -8798,1146  |                    |                             |                    |                                   |                    |
| хи-квадрат   | 2389,66   |                    |                             |                    |                                   |                    |
| Тест отношения правдоподобия                                   | $H_0: \rho_{21} = \rho_{31} = \rho_{32} = 0,$<br>chi2(3)= 50,1122, p-value = 0,0000 |                    |                             |                    |                                   |                    |
| Число наблюдений   | 10 102  |                    |                             |                    |                                   |                    |

\*\*\*, \*\*, \* значим на уровне значимости 1, 5 и 10% соответственно. В скобках указаны эталонные категории.



На наличие болезней сердца и позвоночника статистически значимое влияние оказывает тип поселения: по сравнению с проживающими в областных центрах иметь заболевания сердца, скорее, будут респонденты, проживающие в поселках городского типа (ПГТ) и в сельской местности, а заболевания позвоночника – респонденты, проживающие в городе и в сельской местности.

Работающие индивиды имеют меньшие шансы страдать заболеваниями сердца. Иметь заболевания легких с меньшей вероятностью будут респонденты с детьми и получающие пенсию. Вполне логично, что в старших возрастных группах (старше 45 лет), скорее будут отмечаться заболевания сердца, легких или позвоночника. При этом в группу риска по заболеваниям легких также попали индивиды в возрасте 35-39 лет.

Ожидаемым оказалось влияние наличия инвалидности: шансы иметь заболевания сердца и легких оказались выше у тех, кто находится на инвалидности. Также очевидно отрицательное влияние хорошей самооценки здоровья на вероятность наличия

у индивида рассматриваемых заболеваний. Однако здесь возможна и обратная ситуация, например, «респондент имеет заболевание (1), что повышает его шансы быть на инвалидности и снижает шансы характеризовать свое здоровье как хорошее».

На наличие заболевания легких у индивида статистически значимо влияют доходы. По сравнению с наиболее низкодоходной группой шансы иметь заболевания легких будут ниже для респондентов, чьи доходы попадают во вторую и четвертую квинтильные группы. Индивиды из наиболее высокодоходной группы имеют больше шансов страдать заболеваниями позвоночника.

На следующем этапе были оценены прогнозные значения вероятностей наступления заболеваний (табл. 2). Средние маргинальные эффекты позволили оценить вероятность наступления каждого отдельного заболевания, а средние прогнозные значения совместной вероятности – вероятности того, что у индивида (1) имеются все три заболевания и (2) все три заболевания отсутствуют.

Таблица 2 – Прогнозы по результатам оценивания многомерной пробит-модели

| № | Прогноз  | Заболевание   | Средняя вероятность           |
|---|--|---|-------------------------------|
| 1 | Частная вероятность наступления одного заболевания | (1) сердца:<br>(2) легких:<br>(3) позвоночника:                           | 0,11255<br>0,06375<br>0,15853 |
| 2 | Совместная вероятность наступления заболеваний     | (1) ни одно заболевание не наступит:<br>(2) все три заболевания наступят: | 0,72993<br>0,00651            |

В среднем вероятность того, что индивид одновременно будет страдать всеми тремя заболеваниями, составила 0,65%, тогда как оценка вероятности того, что у него все три заболевания отсутствуют, составила около 73%. Отметим, что оценка вероятности возникновения одновременно трех заболеваний традиционным актуарным способом, а именно, как произведение вероятностей (точнее, их оценок – долей болеющих), составила 0,114%, что существенно отличается от оценки, вычисленной по результатам оценивания многомерной пробит-модели.

#### Заключение

В современных российских условиях корректная оценка вероятности наступле-

ния страхового события осложняется отсутствием достоверных релевантных статистических данных, что влияет на эффективность и устойчивость страховой деятельности. В связи с этим приоритетным направлением актуарных исследований является разработка методик и подходов получения наиболее адекватных реальности оценок вероятностей на основе доступной информации. Проведенные расчеты показали, что современные многомерные пробит-модели (*multivariate probit*) представляют собой подходящий для этого инструментарий. Использование данных о факторах появления у индивида определенных заболеваний дает возможность по-

лучить более точный прогноз того, заболит ли застрахованный в течение действия договора страхования каким-то одним из них или всеми одновременно, по сравнению со стандартной актуарной методикой, основанной на теоремах теории вероятностей. В связи с тем что у многомерных пробит-моделей отсутствует ограничение на число уравнений в модели, необходимые вычисления можно провести для любого количества заболеваний, включенных в покрытие страхового полиса.

Перспективы применения многомерных пробит-моделей для целей актуарных расчетов видятся в двух направлениях. Первое связано с возможностью использования при формировании страховых тарифов средних вероятностей наступления заболеваний, полученных по результатам оценивания данных моделей. Второе направлено на формирование системы поправочных коэффициентов к уже рассчитанным тарифам, позволяющих повышать или снижать тариф на основе информации о значениях индивидуальных характеристик потенциального застрахованного. Это может быть реализовано на основе вычисленных по модели прогнозных значений вероятностей, при условии конкретных значений отдельных характеристик индивида. Например, при страховании на случай возникновения заболеваний легких корректировочный коэффициент, и, соответственно, страховой тариф будет выше для членов низкодоходных групп, а для заболеваний позвоночника – напротив, высокодоходных. Возможность учета одновременности возникновения комплекса заболеваний создает дополнительные преимущества многомерных пробит-моделей и стимулирует продолжение исследований в данном направлении.

#### **БИБЛИОГРАФИЧЕСКИЙ СПИСОК**

1. Жигалкин, М.Ю. Разработка системы тарификации при наличии категориальных рейтинговых факторов / М.Ю. Жигалкин// *Страховое дело*. – 2005. – №3. – С. 39-49.
2. Кудрявцев, А.А., Плам, Р.Г., Чернова, Г.В. *Страхование здоровья (опыт Великобритании)* / А.А. Кудрявцев, Р.Г. Плам, Г.В. Чернова. – М.: Анкил, 2003. – 216 с.
3. Синявская Т.Г., Трегубова А.А. *Возможность корректировки тарифов страхования жизни с учетом регионального риска/ Т.Г. Синявская, А.А. Трегубова// Математическое моделирование в экономике и управлении рисками : материалы III Междунар. молодежной науч.-практ. конф. – Саратов: Изд-во Саратов. ун-та, 2014. – С.340-346.*
4. Синявская Т.Г., Трегубова А.А. *Многомерные пробит-модели как инструмент оценки взаимосвязанных рисков в добровольном медицинском страховании/ Т.Г. Синявская, А.А. Трегубова// Математическое моделирование в экономике и управлении рисками: материалы IV Междунар. молодежной науч.-практ. конф.: в 2 т./ Т.1: Математическое и компьютерное моделирование экономических процессов. – Саратов: Изд-во Саратов. ун-та, 2015. – С.222-228.*
5. Синявская Т.Г., Трегубова А.А. *Проблемы статистической оценки риска в личном страховании: Монография/ Т.Г. Синявская, А.А. Трегубова// – Ростов-на-Дону: Издательско-полиграфический комплекс РГЭУ (РИНХ), 2014. – 125 с.*
6. Синявская Т.Г., Трегубова А.А. *Статистическая оценка взаимосвязанных рисков в медицинском страховании: применение многомерных пробит-моделей/ Т.Г. Синявская, А.А. Трегубова// Мы продолжаем традиции российской статистики: Материалы I Открытого российского статистического конгресса (Новосибирск, 20-22 октября 2015 года). – Новосибирск: НГУ-ЭУ, 2015. – С.53-54.*
7. Синявская Т.Г., Трегубова А.А. *Учет регионального риска в корректировочных коэффициентах к тарифам страхования жизни/ Т.Г. Синявская, А.А. Трегубова// Труды X Международной научной конференции «Применение многомерного статистического анализа в экономике и оценке качества» (26–28 августа 2014 г.). / Национальный исследовательский институт «Высшая школа экономики», Центральный экономико-*

математический институт Российской академии наук. – Москва, 2014. – С. 130-132.

8. Синявская, Т.Г., Трегубова, А.А. Подходы к оценке тарифов в страховании жизни индивидов с повышенным риском (на примере курения) / Т.Г. Синявская, А.А. Трегубова// *Страховое дело*. – 2011. – №5 (220). – С. 27-33.

9. Сухинин, В.Ю., Плаксина, Н.Н. Страхование на случай возникновения смертельно опасных заболеваний / В.Ю. Сухинин, Н.Н. Плаксина// *Страховое дело*. – 1997. – №12. – С. 34-42.

10. Портал «Страхование сегодня». [Электронный ресурс]. Режим доступа: <http://www.insur-info.ru/statistics/analytics/>

11. Сайт обследования RLMS-HSE. Режим доступа: <http://www.hse.ru/rlms/>

12. Ashford, J., and R. Sowden (1970) Multivariate probit analysis, *Biometrics*, 26(3), pp. 535-546.

13. Balia, S., and A.M. Jones (2004). Mortality, lifestyle and socio-economic status. University of York, Working paper, dated October 2004.

14. Cappellari, L., Jenkins, S.P. Calculation of multivariate normal probabilities by simulation, with applications to maximum simulated likelihood estimation. IZA Discussion Paper No. 2112, May 2006. Bonn, Germany. Available at: [repec.iza.org/dp2112.pdf](http://repec.iza.org/dp2112.pdf)

15. Cappellari, L., Jenkins, S.P. Multivariate probit regression using simulated maximum likelihood (2003) // *The Stata Journal* 3 (2003), Number 3, pp. 278-294. Available at: <http://www.stata-journal.com/sjpdf.html?articlenum=st0101>

16. Dash, A., Grimshaw, D. Dread Disease Cover. An Actuarial Perspective. Presented to Staple Inn Actuarial Society, 1990.

17. Gibbons, R.D., and V. Wilcox-Gök (1998) Health service utilization and insurance coverage: a multivariate probit approach. *Journal of the American Statistical Association*, 93(441), pp. 63-72.

18. Valdez, E.A., Frees, E.W. Longitudinal modeling of Singapore motor insurance, University of New South Wales and the University of Wisconsin-Madison, Working Paper, 28 December 2005. Available at:

<http://wwdocs.fce.unsw.edu.au/actuarial/research/papers/2006/Valdez-Frees-2005.pdf>

19. Young, G., Valdez, E.A., Kohn, R. Multivariate probit models for conditional claim-types. 2006. Available at: <http://www.actuaries.org/afir/colloguia/stockholm/young.pdf>

20. Pulmonary Heart Disease (Cor Pulmonale). Medical and health information from AMN. Available at: [http://www.health.am/cardio/more/pulmonary\\_heart\\_disease\\_cor\\_pulmonale](http://www.health.am/cardio/more/pulmonary_heart_disease_cor_pulmonale)

#### **BIBLIOGRAPHIC LIST**

1. Zhigalkin M.Y. Development of tariff system in the presence of categorical rating factors // *Insurance business*, 2005. – №3. – pp. 39-49.

2. Kudryavtsev A.A., Plum R.G., Chernova G.V. Health insurance (UK experience). – Moscow, 2003. – 216 p.

3. Sinyavskaya T.G., Tregubova A.A. Possibility of life insurance tariffs adjusting with the regional risk// *Mathematical Modelling in Economics and Risk Management: III International Youth Research and Practice Conference materials*. – Saratov, 2014. – pp. 340-346.

4. Sinyavskaya T.G., Tregubova A.A. Multivariate probit models as a way to statistical assess dependent risks in health insurance // *Mathematical Modelling in Economics and Risk Management: IV International Youth Research and Practice Conference materials: in 2 vol./ Vol.1: Mathematical and computer modeling of economic processes*. – Saratov, 2015. – pp. 222-228.

5. Sinyavskaya T.G., Tregubova A.A. Problems of statistical risk assessment in personal insurance: Monograph. – Rostov-on-Don, 2014. – 125 p.

6. Sinyavskaya T.G., Tregubova A.A. Statistic evaluation of dependent risks in health insurance: Multivariate probit models application // *1st Open Russian Statistical Congress (1st ORSC) materials*. – Novosibirsk, 2015. – pp. 53-54.

7. Sinyavskaya T.G., Tregubova A.A. Considering the regional risk in life insurance rates calculating // *The application of multivariate statistical analysis in Economics and quality assessment: X International Research*

Conference materials / National Research University "Higher School of Economics". – Moscow, 2014. – pp. 130-132.

8. Sinyavskaya T.G., Tregubova A.A. Approaches to premium estimation in life insurance for individuals with the raised risk (on a smoking example)// Insurance business, 2011. – №5 (220). – pp. 27-32.

9. Sukhinin W.Y., Plaksina N.N. Critical Illness Insurance// Insurance business, 1997. – №12. – pp. 34-42.

10. JSC MIG "Insurance Today". Available at: <http://www.insur-info.ru/statistics/analytics/>

11. The Russia Longitudinal Monitoring Survey – Higher School of Economics (RLMS-HSE). Available at: <http://www.hse.ru/rlms/>

12. Ashford, J., and R. Sowden (1970) Multivariate probit analysis, *Biometrics*, 26(3), pp. 535-546.

13. Balia, S., and A.M. Jones (2004). Mortality, lifestyle and socio-economic status. University of York, Working paper, dated October 2004.

14. Cappellari, L., Jenkins, S.P. Calculation of multivariate normal probabilities by simulation, with applications to maximum simulated likelihood estimation. IZA Discussion Paper No. 2112, May 2006. Bonn, Germany. Available at: [repec.iza.org/dp2112.pdf](http://repec.iza.org/dp2112.pdf)

15. Cappellari, L., Jenkins, S.P. Multivariate probit regression using simulated maximum likelihood (2003) // *The Stata Journal* 3 (2003), Number 3, pp. 278-294. Available at: <http://www.stata-journal.com/sjpdf.html?articlenum=st0101>

16. Dash, A., Grimshaw, D. Dread Disease Cover. An Actuarial Perspective. Presented to Staple Inn Actuarial Society, 1990.

17. Gibbons, R.D., and V. Wilcox-Gök (1998) Health service utilization and insurance coverage: a multivariate probit approach. *Journal of the American Statistical Association*, 93(441), pp. 63-72.

18. Valdez, E.A., Frees, E.W. Longitudinal modeling of Singapore motor insurance, University of New South Wales and the University of Wisconsin-Madison, Working Paper, 28 December 2005. Available at: <http://wwwdocs.fce.unsw.edu.au/actuarial/research/papers/2006/Valdez-Frees-2005.pdf>

19. Young, G., Valdez, E.A., Kohn, R. Multivariate probit models for conditional claim-types. 2006. Available at: <http://www.actuaries.org/afir/colloguia/stockholm/young.pdf>

20. Pulmonary Heart Disease (Cor Pulmonale). Medical and health information from AMN. Available at: [http://www.health.am/cardio/more/pulmonary\\_heart\\_disease\\_cor\\_pulmonale](http://www.health.am/cardio/more/pulmonary_heart_disease_cor_pulmonale)