

**UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY**

Ekonomická a finančná matematika



**Modelovanie vzťahov na trhu práce pomocou
metód maximálnej vierohodnosti**

DIPLOMOVÁ PRÁCA

Michal Gemeran

Bratislava, 2008

Modelovanie vzťahov na trhu práce pomocou metód maximálnej vierohodnosti

DIPLOMOVÁ PRÁCA

Michal Gemeran

**UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY**

Katedra aplikovanej matematiky a štatistiky

Ekonomická a finančná matematika

Ing. Vladimír Kvetan

Bratislava, 2008

Čestne prehlasujem, že som diplomovú prácu vypracoval samostatne, s využitím teoretických vedomostí a s použitím uvedenej literatúry.

V Bratislave 28. apríla 2008

.....
Michal Gemeran

Ďakujem vedúcemu diplomovej práce Ing. Vladimírovi Kvetanovi za trpezlivosť, cenné rady, pripomienky a odborné vedenie práce. Týmto by som sa chcel poďakovať aj Mgr. Ing. Michalovi Páleníkovi za ochotu a pomoc pri poskytnutí dát k tejto diplomovej práci. Ďakujem aj svojim blízkym, ktorí mi boli oporou počas písania diplomovej práce.

Abstrakt

Štruktúra trhu práce v Slovenskej republike je výrazne ovplyvňovaná regionálnou diferencovanosťou, nízkou flexibilitou a mobilitou pracovnej sily, vzdelanosťou a vekovou štruktúrou obyvateľstva a nízkou mierou aktivity uchádzačov o pracovné miesto, o čom svedčí aj vysoká miera prirodzenej nezamestnanosti. Cieľom tejto diplomovej práce je pokúsiť sa popísať vplyv kvalitatívnych znakov na úspešnosť na trhu práce. Pri modelovaní použijeme metódy maximálnej vierohodnosti – konkrétne modely Logit a Probit. Budeme sa venovať aj metodike zisťovania informácií o stave trhu práce na Slovensku, predovšetkým Výberovému zisťovaniu pracovných síl. Na záver sa pokúsime vypočítať kvantitatívny vplyv modelovaných kvalitatívnych znakov na úspešnosť na trhu práce.

Kľúčové slová: Logit model, Probit model, trh práce, nezamestnanosť, metódy maximálnej vierohodnosti, Výberové zisťovanie pracovných síl

Abstract

Labour market structure in Slovak Republic is strongly influenced by regional differentiation, low flexibility and mobility of labour force, education and age structure of the population and low rate of activity of job applicants. The evidence of this fact gives also the high level of Non-accelerating inflation rate of unemployment. The objective of this master thesis is an attempt to define the influence of the qualitative attributes on the labour market successfulness. Maximum likelihood methods, specifically Logit and Probit models, will be used in our simulation. We'll also discuss the methods of collecting information about labour market in Slovakia, mainly Labour force sample survey. At the end we will try to calculate qualitative influence of simulated quantitative attributes on the labour market successfulness.

Keywords: Logit model, Probit model, labour market, unemployment, Maximum likelihood methods, Labour force sample survey

Obsah

Úvod.....	8
1 Regresná analýza.....	10
1.1 Regresná rovnica.....	10
1.2 Lineárny regresný model (LRM)	10
1.2.1 Definícia modelu	10
1.2.2 Základné predpoklady modelu	11
1.2.3 Odhad a interpretácia parametrov v LRM.....	12
1.3 Lineárny pravdepodobnostný model (LPM)	13
1.3.1 Regresia s binárnou závislou premennou	13
1.3.2 Definícia modelu	13
1.3.3 Interpretácia parametrov v LPM	14
1.3.4 Problémy pri odhadovaní parametrov v LPM	14
1.4 Logit a Probit model	16
1.4.1 Definícia modelov.....	16
1.4.2 Logit vs. Probit model.....	17
1.4.3 Interpretácia odhadov parametrov v Logit a Probit modeli	18
1.4.4 Zovšeobecnený Logit a Probit model	19
1.4.5 Interpretácia parametrov v zovšeobecnenom Logit a Probit modeli	20
1.4.6 Metóda latentnej premennej	21
1.4.7 Odhad parametrov v Logit a Probit modeloch	22
1.4.8 Testovanie významnosti modelov	23
1.4.9 Miera priliehavosti v Logit a Probit modeloch	25
1.4.10 Miera správnosti predikcie	26
2 Trh práce na Slovensku.....	27
2.1 Historický vývoj.....	27
2.2 Problémy trhu práce na Slovensku	29
2.3 Zber informácií o trhu práce na Slovensku	37
2.3.1 Metodiky zberu informácií o trhu práce	37
2.3.2 Výberové zisťovanie pracovných síl (VZPS)	38
3 Modelovanie úspešnosti na trhu práce	40
3.1 Príprava modelovania.....	40
3.1.1 Výber individuálnych dát.....	40
3.1.2 Voľba premenných modelu a formulácia hypotéz	41
3.1.3 Popis kvalitatívnych premenných.....	43
3.2 Formulácia všeobecného modelu.....	44
3.2.1 Všeobecný model	44

3.2.2	Logit model.....	44
3.2.3	Probit model.....	44
4	<i>Analýza výsledkov.....</i>	45
4.1	Úvod	45
4.2	Výsledky odhadov parametrov.....	45
4.3	Analýza priliehavosti modelov	49
	<i>Záver</i>	51
5	<i>Použitá literatúra a pramene</i>	53
5.1	Použitá literatúra	53
5.2	Pramene.....	54
6	<i>Prílohy</i>	56
6.1	Odvodenie rovnice pre Logit model	56
6.2	Odvodenie hraničného vplyvu pre Logit model	56
6.3	Tabuľky odhadnutých modelov	57
6.4	Krížové tabuľky	61

Zoznam grafov

Graf 1.1:	Porovnanie distribučnej funkcie Logit a Probit modelu	18
Graf 2.1:	Vývoj miery nezamestnanosti v SR	28
Graf 2.2:	Vývoj počtu nezamestnaných v SR	28
Graf 2.3:	Vývoj počtu nezamestnaných v SR podľa vekových skupín	31
Graf 2.4:	Pomer počtu nezamestnaných v SR podľa vekových skupín.....	31
Graf 2.5:	Vývoj počtu nezamestnaných v SR podľa vzdelania.....	33
Graf 2.6:	Pomer počtu nezamestnaných v SR podľa vzdelania	33
Graf 2.7:	Vývoj počtu nezamestnaných v SR podľa dĺžky nezamestnanosti	34
Graf 2.8:	Pomer nezamestnaných v SR podľa dĺžky nezamestnanosti.....	35
Graf 2.9:	Vývoj počtu nezamestnaných mužov a žien v SR	36
Graf 2.10:	Pomer nezamestnaných mužov a žien v SR	36
Graf 2.11:	Porovnanie vývoja miery nezamestnanosti v SR.....	37

Zoznam tabuliek

Tabuľka 4.1: Odhady parametrov a ich vzájomný vzťah v Logit a Probit modeli	45
Tabuľka 4.2: P-hodnoty odhadnutých parametrov v jednotlivých modeloch	46
Tabuľka 4.3: Hraničné vplyvy odhadnutých parametrov v jednotlivých modeloch	48
Tabuľka 4.4: Miery priliehavosti modelov	49
Tabuľka 4.5: Predikcia statusu zamestnanosti v jednotlivých modeloch	49
Tabuľka 6.1: Logit model za rok 2000	57
Tabuľka 6.2: Probit model za rok 2000	58
Tabuľka 6.3: Logit model za rok 2006	59
Tabuľka 6.4: Probit model za rok 2006	60
Tabuľka 6.5: Krížová tabuľka premenných ILO a SEX pre rok 2000	61
Tabuľka 6.6: Krížová tabuľka premenných ILO a YOUNG pre rok 2000	62
Tabuľka 6.7: Krížová tabuľka premenných ILO a OLD pre rok 2000	63
Tabuľka 6.8: Krížová tabuľka premenných ILO a EDU pre rok 2000	64
Tabuľka 6.9: Krížová tabuľka premenných ILO a C_C pre rok 2000	65
Tabuľka 6.10: Krížová tabuľka premenných ILO a SINGLE pre rok 2000	66
Tabuľka 6.11: Krížová tabuľka premenných ILO a MARRIED pre rok 2000	67
Tabuľka 6.12: Krížová tabuľka premenných ILO a DIVORCED pre rok 2000	68
Tabuľka 6.13: Krížová tabuľka premenných ILO a SEX pre rok 2006	69
Tabuľka 6.14: Krížová tabuľka premenných ILO a YOUNG pre rok 2006	70
Tabuľka 6.15: Krížová tabuľka premenných ILO a OLD pre rok 2006	71
Tabuľka 6.16: Krížová tabuľka premenných ILO a EDU pre rok 2006	72
Tabuľka 6.17: Krížová tabuľka premenných ILO a C_C pre rok 2006	73
Tabuľka 6.18: Krížová tabuľka premenných ILO a SINGLE pre rok 2006	74
Tabuľka 6.19: Krížová tabuľka premenných ILO a MARRIED pre rok 2006	75
Tabuľka 6.20: Krížová tabuľka premenných ILO a DIVORCED pre rok 2006	76

Úvod

Úspešnosť na trhu práce predstavuje skutočnosť, či si človek ako jedinec s individuálnymi vlastnosťami, dokáže nájsť prácu, alebo sa stane nezamestnaným. Faktory, ktorými je „obdarený“ každý z nás, ovplyvňujú potenciálny úspech pri hľadaní zamestnania v rozličnej miere. V rôznorodých slovenských podmienkach môže byť veľmi zaujímavé pozorovať tento sociálno-ekonomický jav. Skúmaniu úspešnosti na trhu práce sa na Slovensku venuje málo pozornosti. Väčšinou sa k nám dostávajú len informácie o dôsledkoch stavu na trhu práce, ale na príčiny a ich riešenie sa kladie menší dôraz. Výsledky tejto diplomovej práce by preto mohli byť obohacujúce nielen pre samotného uchádzača o zamestnanie, ale aj pre rôzne sociálne inštitúcie zaoberajúce sa touto dôležitou problematikou. Navyše otázka nezamestnanosti rezonuje v odborných kruhoch či medzi laickou verejnosťou od vzniku Slovenskej republiky až po súčasnosť. Z národohospodárskeho hľadiska je nezamestnanosť tiež dôležitým ukazovateľom hodnotenia výkonnosti ekonomiky. Aj keď má miera nezamestnanosti v ostatnom čase na Slovensku klesajúcu tendenciu, štruktúra trhu práce, najmä čo sa týka zloženia dlhodobo nezamestnaných, sa výrazne nezmenila. Žiadnej vláde sa doposiaľ nepodarilo spôsobiť v nej zásadnú zmenu. Výrazný vplyv na tom má najmä laxný prístup k aktívnej politike zamestnanosti.

Otázka nezamestnanosti je pre Slovensko, ako člena Európskej únie, dvojnásobne dôležitá. Z pohľadu Európskej únie sa viaceré krajiny v súčasnosti aktívne zameriavajú na zvýšenie flexibility pracovných trhov, zníženie štrukturálnej nezamestnanosti a zjednodušenie prechodu k otvorenejšiemu trhu práce. Z výskumov Európskej komisie vyplýva, že pre mnohé členské krajiny je dlhodobým problémom nezamestnanosť mladých ľudí. V porovnaní s celkovým priemerom pracujúcej populácie je u mladých viac ako dvojnásobná pravdepodobnosť, že budú nezamestnaní. Hoci pracuje čoraz viac žien a starších ľudí, nezamestnanosť mladých ľudí sa udržiava na vysokej hodnote 17,4 %. Mladí Európania často nemajú postačujúce vzdelanie. Takmer každý šiesty mladý človek skončí školskú dochádzku priskoro. EÚ má v úmysle tento podiel znížiť, a preto odporúča zvýšenie investícií do vzdelávania a odbornej prípravy, aby sa zvýšila úroveň dosiahnutého vzdelania.¹

¹ Podľa správ Európskej komisie <http://ec.europa.eu>.

Cieľom tejto diplomovej práce je pokúsiť sa modelovať vplyv kvalitatívnych znakov na úspešnosť na trhu práce. Metódy maximálnej vierohodnosti nám poslúžia na špecifikovanie tohto vzťahu. Samotná diplomová práca sa skladá z viacerých častí.

Úvodná časť diplomovej práce sa zaoberá teoretickou prípravou modelovania. Postupne sa venuje popisu modelovacích techník od základných predpokladov jednoduchého lineárneho regresného modelu až k podrobnému popisu modelov s kvalitatívnou závislou premennou - Logit a Probit. Tieto regresné modely sú veľmi užitočné napr. pri analýzach prierezných alebo panelových dát. Údaje pre takéto analýzy pochádzajú z rôznych prieskumov verejnej mienky, v ktorých človek vyjadruje spôsob svojej voľby formou, ktorá sa dá popísať práve kvalitatívnou premennou (napr. či bude voliť v referende áno alebo nie, či je zamestnaný alebo nie a pod.). Logit a Probit modelom sa medzi odbornou verejnosťou venuje relatívne málo pozornosti, o čom svedčí aj fakt, že nie sú zahrnuté vo vyučovacom procese na FMFI UK. To je ďalší z dôvodov, prečo je v diplomovej práci tejto skupine regresných modelov venovaný dostatočne široký priestor. Pri modelovaní úspešnosti na trhu práce sa zameriavame aj na sledovanie podobnosti Logit a Probit modelov. V závere úvodnej časti sú popísané metódy odhadov parametrov spomenutých modelov a miery kvality modelovania vzájomných vzťahov medzi jednotlivými premennými.

Druhá časť práce sa zaoberá stavom trhu práce na Slovensku z pohľadu vývoja nezamestnanosti. Historický vývoj nezamestnanosti naznačuje pretrvávajúcu existenciu problémov v štruktúre trhu práce na Slovensku. Podrobnejšiemu popisu konkrétnych problémov je venovaná celá kapitola. Od popisu stavu na trhu práce sú neskôr odvodené nezávislé premenné do modelových rovníc. Záver tejto časti je venovaný popisu metodiky zberu dát o trhu práce - Výberovému zisťovaniu pracovných síl. Dáta použité pri modelovaní pochádzajú práve z tohto zdroja.

Tretia časť diplomovej práce sa venuje príprave modelovania z hľadiska výberu a klasifikácie premenných do modelu, formulovaniu hypotéz o očakávaných výsledkoch a samotnému modelovaniu úspešnosti na trhu práce.

V poslednej časti sú analyzované a zhrnuté výsledky modelovania a overujú sa tiež vyslovené hypotézy o správaní trhu práce na Slovensku.

1 Regresná analýza

1.1 Regresná rovnica

Regresné modely sa používajú na špecifikáciu funkčných vzťahov a analyzovanie závislosti medzi jednou vysvetľovanou (endogénnou) premennou a jednou alebo viacerými vysvetľujúcimi (exogénnymi) premennými. Vysvetľovanú premennú môžeme nazvať aj náhodnou, odhadovanou, závislou alebo premennou odozvy. Vysvetľujúce premenné môžeme pomenovať nenáhodnými, pozorovanými, nezávislými premennými alebo regresormi. V regresnej rovnici vystupuje závislá premenná ako funkcia nezávislých premenných (s prislúchajúcimi neznámymi parametrami) a náhodnej zložky ².

Označme si náhodnú premennú y , pozorované veličiny $X_1, X_2, \dots, X_k$, a náhodnú zložku ε . Potom môžeme zapísať všeobecný funkčný vzťah pre regresnú rovnicu:

$$y = f(X_1, X_2, \dots, X_k, \varepsilon)$$

1.2 Lineárny regresný model (LRM) ³

1.2.1 Definícia modelu

V *lineárnom regresnom modeli* sa pri modelovaní snažíme odhadnúť *lineárny* vzťah medzi vysvetľujúcimi premennými a vysvetľovanou premennou. Lineárna regresia sa snaží popísať tento funkčný vzťah ako *priamku* priliehajúcu k pozorovaným dátam. Závislá premenná predstavuje lineárnu kombináciu parametrov, nemusí však byť lineárna v nezávislých premenných.

Lineárny regresný model a jeho modifikácie sa v praxi často používajú v rôznych vedeckých odvetviach, napr. v biológii, medicíne, ekonómii, štatistike atď.

² Viac v Hatrák (2007) [2.].

³ Inšpirácia čerpaná z Hatrák 2007 [2.], Klimešová 2007 [4.], Witkovský 2007 [19.].

Predpokladajme, že máme n -rozmernú modelovanú premennú $\mathbf{y}$ ⁴ a chceme odhadnúť jej závislosť na nezávislých premenných $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k$. Nech $\beta_0, \beta_1, \dots, \beta_k$ ⁵ sú konštantné parametre prislúchajúce k nezávislým premenným. Tento funkčný vzťah môžeme zapísať pre i -te pozorovanie premennej $\mathbf{y}$ takto:

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i$$

alebo v nasledovnom maticovom tvare:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

kde:

$$\mathbf{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

$\mathbf{y}$ - n -rozmerný vektor vysvetľovaných veličín

$\mathbf{X}$ - $n \times (k + 1)$ dátová matica vysvetľujúcich premenných

$\boldsymbol{\beta}$ - $(k + 1)$ -rozmerný vektor neznámych parametrov modelu

$\boldsymbol{\varepsilon}$ - n -rozmerný vektor neznámych náhodných chýb

1.2.2 Základné predpoklady modelu⁶

1. Platnosť modelu (funkčnej formy): $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$
2. Identifikačná podmienka: $h(\mathbf{X}) = k + 1$, matica $\mathbf{X}$ typu $n \times (k + 1)$ má plnú hodnotu a $n > (k + 1)$, teda počet pozorovaní je väčší ako počet neznámych parametrov modelu
3. Pre strednú hodnotu náhodných chýb platí: $E(\varepsilon_i) = 0$, sú rozmiestnené rovnomerne
4. Pre varianciu náhodných chýb platí: $\text{var}(\varepsilon_i) = \sigma^2$, čo znamená požiadavku homoskedasticity, pričom σ^2 je neznámy parameter. Ďalej musí platiť $\text{cov}(\varepsilon_i, \varepsilon_j) = 0$, pre $i \neq j, i, j = 1, \dots, n$, jedná sa o párovú nekorelovanosť chýb

⁴ Nahrubo (Bold) zvýraznené premenné predstavujú matice, resp. vektory.

⁵ Medzi neznámymi parametrami modelu môže vystupovať aj konštantný parameter β_0 , ktorý predstavuje priesečník priamky definovanej regresnou rovnicou s horizontálnou osou vysvetľujúcich premenných.

⁶ Podmienky 1-5 sa niekedy nazývajú aj Gauss-Markovove podmienky (GM podmienky).

5. Nekorelovanosť chýb a vysvetľujúcich premenných: $Cov(\mathbf{X}, \boldsymbol{\varepsilon}) = E(\mathbf{X}'\boldsymbol{\varepsilon}) = 0$
6. Normalita rozdelenia chýb: $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$

Na základe vyššie uvedených predpokladov môžeme napísať:

$$E(\mathbf{y}) = E(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) = E(\mathbf{X}\boldsymbol{\beta}) + E(\boldsymbol{\varepsilon}) = \mathbf{X}\boldsymbol{\beta}$$

$$Var(\mathbf{y}) = Var(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) = Var(\mathbf{X}\boldsymbol{\beta}) + Var(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}$$

kde $\mathbf{I}$ je jednotková matica typu $n \times n$.

1.2.3 Odhad a interpretácia parametrov v LRM

Najčastejšou metódou, používanou pri odhadovaní parametrov v LRM, je metóda najmenších štvorcov⁷. Po aplikácii tejto metódy na lineárnu regresiu dostávame nasledujúci odhadnutý vzťah:

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_k X_{ik} + e_i$$

kde e_i predstavuje reziduál, definovaný ako $e_i = y_i - \hat{y}_i$ a $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k, \hat{y}_i$ predstavujú odhadnuté hodnoty pre $\beta_0, \beta_1, \dots, \beta_k, y_i$.

Hraničný vplyv vysvetľujúcej premennej X_{ij} na predpovedanú vysvetľovanú premennú $\hat{y}_i$ dostaneme derivovaním:

$$\frac{\partial \hat{y}_i}{\partial X_{ij}} = \frac{\partial (\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_k X_{ik})}{\partial X_{ij}} = \hat{\beta}_j$$

Z čoho môžeme vidieť, že ak zväčšíme vysvetľujúcu premennú X_{ij} o jednotku a ostatné premenné ponecháme nezmenené, odhadnutý parameter $\hat{\beta}_j$ prislúchajúci k tejto premennej vyjadruje veľkosť a smer zmeny v predpovedanej hodnote vysvetľovanej veličiny $\hat{y}_i$.

⁷ Spomenutej metóde sa tu nebudeme podrobnejšie venovať, popísaná je napr. v Hatrák 2007 [2.].

1.3 Lineárny pravdepodobnostný model (LPM) ⁸

1.3.1 Regresia s binárnou závislou premennou

Regresný model s *binárnou*⁹ závislou premennou popisuje situáciu voľby medzi dvoma alternatívami v závislosti od známych skutočností, napr. či sa človek rozhodne pre kúpu auta, ak má mesačný príjem v určitej výške a pod. V nasledujúcej časti budeme diskutovať o tom, či môžeme pri odhadovaní takéhoto modelu použiť regresiu a ako sa v takomto prípade zmení metóda odhadovania neznámych parametrov a ich interpretácia.

1.3.2 Definícia modelu

Predpokladajme, že v modeli vystupuje závislá *binárna* premenná y , ktorá nadobúda len dve možné hodnoty. V prípade, že udalosť nastala, má hodnotu 1 a inak hodnotu 0. Vektor nezávislých premenných X môže obsahovať kvantitatívne premenné, transformácie kvantitatívnych premenných, polynomiálne regresory sformované z kvantitatívnych premenných, binárne regresory reprezentujúce kvalitatívne premenné alebo aj interaktívne regresory.

Pre jednoduchosť sa obmedzme na regresný model pozostávajúci len z jednej vysvetľujúcej premennej:

$$y_i = \beta X_i + \varepsilon_i.$$

Nech y_i má alternatívne rozdelenie s pravdepodobnosťou výskytu udalosti π_i :

- $P(y_i = 1) = \pi_i$ ¹⁰
- $P(y_i = 0) = 1 - \pi_i$, pre $i = 1, \dots, n$.

Pre strednú hodnotu a rozptyl takejto diskkrétnej náhodnej premennej y_i platí:

$$E(y_i) = 1 \cdot P(y_i = 1) + 0 \cdot P(y_i = 0) = P(y_i = 1) = \pi_i$$

⁸ Inšpirácia čerpaná zo Stankovičová 2007[11.], Stankovičová, Vojtková 2007[12.], Fox 2007 [15.], Leblang 2004 [16.], Murphy 2007 [17.].

⁹ V literatúre sa často môžeme stretnúť s výrazom „dummy“ premenná, čo je označenie pre veličinu, ktorá má podobu slovných charakteristík pre jednotlivé pozorovania.

¹⁰ Ide o podmienenú pravdepodobnosť $P(y_i = 1 | X_i)$ nastatia javu y_i za podmienky výskytu premennej X_i . V ďalšom texte sa obmedzíme len na výraz $P(y_i = 1)$ a budeme predpokladať, že ide o podmienenú pravdepodobnosť.

Za platnosti predpokladov pre LRM môžeme pre binárnu vysvetľovanú premennú y_i napísať:

$$E(y_i) = \pi_i = \beta X_i$$

V tomto prípade modelujeme pravdepodobnosť úspechu alebo výskytu určitej udalosti, napr. pravdepodobnosti, či si respondent s určitou množinou jemu vlastných atribútov vyberie skôr jednu alternatívu ako druhú. Preto sa takýto model nazýva *lineárny pravdepodobnostný model*.

1.3.3 Interpretácia parametrov v LPM

Interpretácia parametrov v LPM je podobná ako v LRM až na jednu zmenu. Ak zväčšíme vysvetľujúcu premennú X_i o jednotku, parameter $\hat{\beta}$ prislúchajúci k tejto premennej, vyjadruje smer a veľkosť zmeny *pravdepodobnosti* predpovedanej vysvetľovanej veličiny $\hat{y}_i$.

1.3.4 Problémy pri odhadovaní parametrov v LPM

Ak chceme použiť regresnú rovnicu a následne pomocou niektorej modelovacej metódy odhadnúť vzájomnú závislosť zvolených premenných, musíme najprv pretransformovať závislú binárnu premennú na veličinu so spojitými hodnotami, ktorá bude funkciou pravdepodobnosti výskytu udalosti. Pri takomto transformačnom procese sa nám však vynárajú viaceré problémy:

1. Náhodné chyby nie sú normálne rozdelené

Keďže binárna premenná y_i nadobúda len dve hodnoty 0 a 1, náhodné chyby ε_i sú takisto binárne a teda nemajú normálne rozdelenie:

- ak $y_i = 1$, čo nastáva s pravdepodobnosťou π_i , potom $\varepsilon_i = 1 - \beta X_i = 1 - E(y_i) = 1 - \pi_i$
- ak $y_i = 0$, čo nastáva s pravdepodobnosťou $1 - \pi_i$, tak $\varepsilon_i = 0 - \beta X_i = 0 - E(y_i) = -\pi_i$

Normalitu náhodných chýb požadujeme kvôli efektívnosti odhadov parametrov.

2. Disperzia náhodných chýb nie je konštantná

Ak je splnený predpoklad linearity $E(\varepsilon_i) = 0$, disperziu náhodných chýb môžeme vyjadriť nasledovne:

$$\text{var}(\varepsilon_i) = \pi_i(1 - \pi_i)^2 + (1 - \pi_i)(-\pi_i)^2 = \pi_i(1 - \pi_i)$$

Pripomeňme, že $E(y_i) = \pi_i = \beta X_i$. Keďže pravdepodobnosť π_i závisí od hodnoty X_i , disperzie náhodných chýb sú funkciou premennej $\mathbf{X}$ a nie sú konštantné, sú heteroskedastické.

Z rovnice ďalej vyplýva, že pozorovania, pre ktoré je π_i blízko k 0 alebo blízko k 1, budú mať relatívne malé disperzie náhodných chýb, zatiaľ čo pozorovania s π_i blízkou k 1/2 budú mať variancie relatívne najväčšie.

3. Nelinearita a iracionálne predikcie

Predpoklad linearity $E(\varepsilon_i) = 0$ sa dá udržať len pre obmedzený rozsah hodnôt vektora $\mathbf{X}$. Ak je rozsah hodnôt vzorky dostatočne široký, potom všetky hodnoty pravdepodobnosti π nespádajú do intervalu $[0,1]$. Interpretácia hodnoty π ako pravdepodobnosti mimo jednotkového intervalu však potom nemá racionálny základ.

Aby sme zabezpečili, že π bude nadobúdať len hodnoty od 0 po 1, použijeme rastúcu¹¹ funkciu $F(\cdot)$, ktorá zobrazí βX_i do jednotkového intervalu a zároveň zachová lineárnu štruktúru modelu:

$$\pi_i = F(\beta X_i)$$

Pre funkciu $F(\cdot)$ bude ďalej platiť $F(-\infty) = 0$ a $F(\infty) = 1$.

Ľubovoľná kumulatívna pravdepodobnostná distribučná funkcia (CDF) spĺňa takéto podmienky.

¹¹ Pindyck, Rubinfeld 2000 [9.], Siebertová 2005 [18.].

1.4 Logit a Probit model ¹²

1.4.1 Definícia modelov

Funkcia $F(\cdot)$ býva volená ako distribučná funkcia normovaného normálneho rozdelenia:

$$\Phi(z) = \int_{-\infty}^z \phi(Z) dZ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{\{-\frac{1}{2}Z^2\}} dZ$$

alebo ešte častejšie ako distribučná funkcia logistického rozdelenia:

$$\Lambda(z) = \int_{-\infty}^z \lambda(Z) dZ = \frac{e^z}{1 + e^z} \quad ^{13}$$

Pričom $\phi(Z)$ a $\lambda(Z)$ predstavujú funkcie hustoty pravdepodobnosti normovaného normálneho rozdelenia, resp. logistického rozdelenia so strednou hodnotou 0 a disperziou $\pi^2/3$. ¹⁴

Použitím distribučnej funkcie normovaného normálneho rozdelenia $\Phi(\cdot)$, dostávame *lineárny Probit model*:

$$\pi_i = \Phi(\beta X_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta X_i} e^{\{-\frac{1}{2}Z^2\}} dZ$$

Použitím distribučnej funkcie logistického rozdelenia $\Lambda(\cdot)$ dostávame *lineárnu logistickú regresiu* alebo *lineárny Logit model*:

$$\pi_i = \Lambda(\beta X_i) = \int_{-\infty}^{\beta X_i} \lambda(Z) dZ = \frac{1}{1 + e^{-(\beta X_i)}} = \frac{e^{\beta X_i}}{1 + e^{\beta X_i}} \quad ^{15}$$

¹² Inšpirované Kališová 2004 [3.], Klimešová 2007 [4.], McFadden 1984 [8.], Pindyck, Rubinfeld 2000 [9.], Senaj 2005 [10.], Croux 2007 [14.], Fox 2007 [15.], Leblang 2004 [16.], Murphy 2007 [17.], Siebertová 2005 [18.], EViews 2004 [21.]

¹³ Štandardná logistická funkcia, tiež nazývana ako sigmoidná krivka alebo „S-krivka“.

¹⁴ $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{\{-\frac{1}{2}z^2\}}$, $\lambda(z) = \frac{e^{-z}}{(1+e^{-z})^2}$

¹⁵ Odvodenie pozri v prílohe 6.1

Upravením základnej rovnice pre Logit model dostaneme nasledujúci vzťah:

$$\frac{\pi_i}{1 - \pi_i} = e^{(\beta X_i)}$$

Pomer $\pi_i/(1 - \pi_i)$ vyjadruje *šance*¹⁶, že $y_i = 1$, teda výskytu želanej udalosti v pomere k výskytu neželanej udalosti.

Zlogaritmovaním¹⁷ oboch strán rovnice pre šance dostávame:

$$\log \frac{\pi_i}{1 - \pi_i} = \beta X_i$$

čo môžeme vyjadriť aj ako inverznú transformáciu pravdepodobnosti:

$$\Lambda^{-1}(\pi_i) = \log \frac{\pi_i}{1 - \pi_i} = \text{logit}(\pi_i)$$

nazývanú tiež *Logit pravdepodobnosti* π_i , predstavujúci logaritmus šancí, že y_i nadobudne skôr hodnotu 1 ako 0.

1.4.2 Logit vs. Probit model

Logit a Probit modely sú si navzájom veľmi podobné. Distribučná funkcia v Logit modeli má „ťažšie chvosty“¹⁸ v porovnaní s Probit modelom. Distribučná funkcia Probit modelu má strmší sklon, ako môžeme vidieť z grafu 1.1. Obe distribučné funkcie sú skoro lineárne medzi $\pi = 0,2$ a $\pi = 0,8$. (Fox, 2007)

Odhady parametrov v Logit a Probit modeloch sa už líšia výraznejšie. Keďže logaritmické rozdelenie má rozptyl rovný $\pi^2/3$, odhady β pochádzajúce z Logit modelu musia byť prenasobené $3^{1/2}/\pi$, aby mohli byť porovnateľné s odhadmi získanými z Probit modelu, ktorý má rozptyl rovný jednej. (Kališová, 2004)

¹⁶ Termín šancí sa používa napr. aj pri stávkovaní a hazardných hrách.

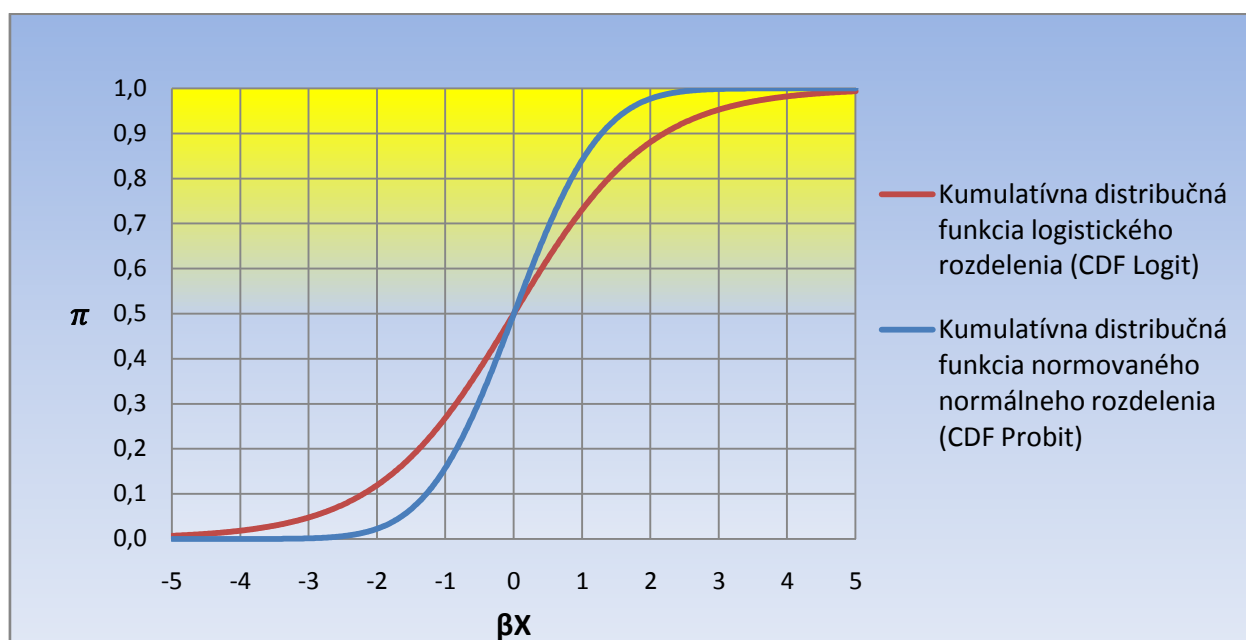
¹⁷ Pri zlogaritmovaní môžeme použiť prirodzený alebo dekadický logaritmus. Výsledky sú analogické, len v praxi sa viac používa prirodzený logaritmus. V ďalšom texte budeme pod označením „log“ myslieť všeobecný symbol pre logaritmus t.j. prirodzený alebo dekadický.

¹⁸ Rozdelenie má viac pozorovaní nachádzajúcich sa na koncoch distribučnej funkcie.

Napriek vzájomnej podobnosti má Logit model pri porovnaní s Probit modelom dve praktické výhody:

1. *Jednoduchosť*: rovnica logistickej distribučnej funkcie je veľmi jednoduchá, zatiaľ čo normálna kumulatívna distribučná funkcia obsahuje nevyčíslený integrál
2. *Interpretovateľnosť*: inverzná lineárna transformácia Logit modelu $\Lambda^{-1}(\pi)$ sa dá priamo interpretovať ako logaritmus šancí, zatiaľ čo inverzná transformácia $\Phi^{-1}(\pi)$ nemá priamu interpretáciu

Graf 1.1: Porovnanie distribučnej funkcie Logit a Probit modelu



1.4.3 Interpretácia odhadov parametrov v Logit a Probit modeli

Odhadnuté koeficienty v Logit a Probit modeloch nemôžeme priamo interpretovať pomocou hraničných vplyvov na vysvetľovanú premennú. Interpretácia odhadnutých parametrov je preto zložitejšia.

Marginálny vplyv na pravdepodobnosť, že $\hat{y}_i = 1$, závisí od hodnoty a znamienka odhadnutého parametra $\hat{\beta}$ a od hodnoty vysvetľujúcej premennej X_i :

$$\frac{\partial P(\hat{y}_i = 1)}{\partial X_i} = \frac{\partial F(\hat{\beta}X_i)}{\partial X_i} = f(\hat{\beta}X_i)\hat{\beta}$$

Kladná hodnota odhadnutého parametra $\hat{\beta}$ znamená, že nárast vo vysvetľujúcej premennej X_i zvýši pravdepodobnosť výskytu premennej odozvy y_i a záporná hodnota $\hat{\beta}$ znamená, že nárast v premennej X_i naopak zníži pravdepodobnosť výskytu premennej odozvy y_i .

Pre Probit model dostávame nasledovný tvar marginálneho vplyvu:

$$\frac{\partial P(\hat{y}_i = 1)}{\partial X_i} = \phi(\hat{\beta}X_i)\hat{\beta}$$

Pre Logit model máme:

$$\frac{\partial P(\hat{y}_i = 1)}{\partial X_i} = \Lambda(1 - \Lambda)\hat{\beta} = \pi(1 - \pi)\hat{\beta}^{19}$$

V prípade, ak $\pi = 1/2$, marginálny efekt môže dosiahnuť svoju maximálnu hodnotu $\hat{\beta}/4$.

Ak upravíme rovnicu pre Logit model na vzťah vyjadrujúci šance, dostávame ďalšiu možnú interpretáciu odhadnutých parametrov:

$$\frac{\pi_i}{1 - \pi_i} = e^{(\hat{\beta}X_i)} = e^{\hat{\beta}X_i} = (e^{\hat{\beta}})^{X_i}$$

Rovnica vyjadruje multiplikatívny efekt pre šance. Ak zväčšíme vysvetľujúcu premennú X_i o jednotku, parameter $\hat{\beta}$ prislúchajúci k tejto premennej bude predstavovať veľkosť zmeny šanci o $e^{\hat{\beta}}$.

1.4.4 Zovšeobecnený Logit a Probit model

Do Logit aj Probit modelu môžeme zahrnúť viaceré vysvetľujúce premenné rôznych typov, podobne ako je to popísané v časti 1.3.2. Dostaneme nasledujúce funkčné vzťahy:

¹⁹ Odvodenie pozri v prílohe 6.2.

Rovnica pre Logit:

$$\pi_i = \Lambda(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik})}}$$

alebo ekvivalentne:

$$\ln \frac{\pi_i}{1 - \pi_i} = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}$$

Podobne pre Probit model:

$$\pi_i = \Phi(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}} e^{\{-\frac{1}{2}Z^2\}} dZ$$

1.4.5 Interpretácia parametrov v zovšeobecnenom Logit a Probit modeli

Interpretácia odhadnutých regresných koeficientov vo všeobecnom Logit a Probit modeli je podobná interpretácii pre Logit a Probit model s jednou vysvetľujúcou premennou.

Marginálny vplyv na pravdepodobnosť, že $\hat{y}_i = 1$, závisí od hodnoty a znamienka odhadnutého parametra $\hat{\beta}_j$, ale aj od hodnoty ostatných odhadnutých parametrov a vysvetľujúcich premenných:

$$\begin{aligned} \frac{\partial P(\hat{y}_i = 1)}{\partial X_{ij}} &= \frac{\partial F(\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_k X_{ik})}{\partial X_{ij}} = \\ &= f(\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_k X_{ik}) \hat{\beta}_j \end{aligned}$$

Uvádzame ešte vplyv odhadnutých parametrov v zovšeobecnenom Logit modeli v zmysle šanci:

$$\frac{\pi_i}{1 - \pi_i} = e^{\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_k X_{ik}} = e^{\hat{\beta}_0} (e^{\hat{\beta}_1})^{X_{i1}} \dots (e^{\hat{\beta}_k})^{X_{ik}}$$

Teda $e^{\hat{\beta}_j}$ predstavuje multiplikatívny efekt na šance pri zvýšení X_{ij} o 1, ak ponecháme ostatné vysvetľujúce premenné konštantné.

1.4.6 Metóda latentnej premennej

Alternatívne odvodenie Logit a Probit modelu je založené na spojitej latentnej vysvetľovanej premennej y^* , ktorá nie je voľne pozorovateľná, a preto ju nemôžeme ani odmerať. Môžeme však pozorovať nasledujúce realizácie premennej y^* :

- $y_i = 0$, ak $y_i^* \leq 0$
- $y_i = 1$, ak $y_i^* > 0$

O latentnej premennej y^* predpokladáme, že je lineárnou funkciou vysvetľujúcej premennej X a nepozorovateľnej chyby ε :

$$y_i^* = \beta X_i + \varepsilon_i$$

Úlohou je odhadnúť parameter β , pričom ale nemôžeme použiť regresnú rovnicu, lebo latentná premenná nie je priamo pozorovaná. Môžeme však vyjadriť pravdepodobnosť latentnej premennej:

$$\pi_i = P(y_i = 1) = P(y_i^* > 0) = P(\beta X_i + \varepsilon_i > 0) = P(\varepsilon_i > -\beta X_i)$$

Môžeme povedať, že $y_i = 1$, ak náhodná zložka je menšia ako systematická časť.

Ak sú chyby rozdelené nezávisle, podľa normovaného normálneho rozdelenia, t.j. $\varepsilon_i \sim N(0,1)$, dostávame Probit model:

$$\pi_i = P(\varepsilon_i > -\beta X_i) = 1 - \Phi(\beta X_i) = \Phi(\beta X_i) \text{ }^{20}$$

Alternatívne, ak ε_i majú logistické rozdelenie, dostávame formuláciu Logit modelu:

$$\pi_i = \Lambda(\beta X_i)$$

Analogické vzťahy dostaneme aj pre zovšeobecnený Logit a Probit model.

²⁰ Posledná rovnosť vyplýva zo symetrickosti rozdelenia.

1.4.7 Odhad parametrov v Logit a Probit modeloch

Na odhad regresných koeficientov v modeloch použijeme metódu maximálnej vierohodnosti, lebo dosahuje konzistentnosť, asymptotickú efektívnosť a normalitu odhadnutých parametrov modelu²¹. Funkcia vierohodnosti nám tiež poslúži pri testovaní významnosti celkového modelu. Odhady parametrov metódou maximálnej vierohodnosti vypočítame maximalizovaním vierohodnostnej funkcie L , resp. jej logaritmu $\log L$ ²², podľa neznámych parametrov.

Predpokladajme, že možnosť $y_i = 1$ nastala vo vzorke práve m_1 - krát a možnosť $y_i = 0$ nastala práve m_2 - krát. Pričom máme $m_1 + m_2 = n$. Vierohodnostná funkcia je funkciou neznámych parametrov a má nasledovný tvar:

$$\begin{aligned} L(\beta) &= \prod_{i=1}^{m_1} P(y_i = 1) \prod_{i=m_1+1}^n P(y_i = 0) \quad {}^{23} = \prod_{i=1}^n P_i^{y_i} (1 - P_i)^{(1-y_i)} = \\ &= \prod_{i=1}^n [F(\beta X_i)]^{y_i} [1 - F(\beta X_i)]^{1-y_i} \end{aligned}$$

Logaritmus funkcie vierohodnosti:

$$\log L(\beta) = \sum_{i=1}^n [y_i \log F(\beta X_i) + (1 - y_i) \log(1 - F(\beta X_i))]$$

Odhad parametra β získame maximalizovaním logaritmu vierohodnostnej funkcie podľa β :

$$\hat{\beta} = \operatorname{argmax}_{\beta} \log L(\beta) = \operatorname{argmax}_{\beta} \log \sum_{i=1}^n [y_i \log F(\beta X_i) + (1 - y_i) \log(1 - F(\beta X_i))]$$

Výsledný odhadnutý parameter $\hat{\beta}$ je ten, ktorý maximalizuje hodnotu vierohodnostnej funkcie:

$$L_{max} = L(\hat{\beta})$$

²¹ Pindyck, Rubinfeld 2000 [9.], Witkovský 2007 [20.].

²² Kvôli zjednodušeniu výpočtov môžeme spraviť logaritmickú transformáciu, ktorá je monotónna a rastúca.

²³ Rovnosť platí, lebo pre prvých m_1 pozorovaní $y_i = 1$ a pre zvyšných m_2 pozorovaní $y_i = 0$.

Logaritmus vierohodnostnej funkcie pre Logit model má tvar:

$$\log L(\beta) = \sum_{i=1}^n [y_i \log \Lambda(\beta X_i) + (1 - y_i) \log(1 - \Lambda(\beta X_i))]$$

Podobne pre Probit model:

$$\log L(\beta) = \sum_{i=1}^n [y_i \log \Phi(\beta X_i) + (1 - y_i) \log(1 - \Phi(\beta X_i))]$$

Na diferencovanie takýchto nelineárnych rovníc musíme použiť numerické metódy.

Analogické vzťahy dostaneme aj pri zovšeobecnenom Logit a Probit modeli.

1.4.8 Testovanie významnosti modelov

Štatistickú významnosť modelov môžeme testovať dvoma spôsobmi.

1. V prvom prípade porovnávame zvolený (úplný) model s modelom s nulovými parametrami, ktorý obsahuje len konštantný koeficient β_0 . Hypotézy pre testovanie sú formulované nasledovne:

- H_0 : model ako celok nie je významný, teda má všetky parametre β_i rovné nule
- H_1 : model ako celok je významný, teda nemá všetky parametre β_i rovné nule

kde $i = 1, \dots, k$ je počet parametrov modelu. Pri testovaní sa využíva vierohodnostná funkcia L (L -štatistika), resp. jej upravený logaritmus $-2 \log L$ ²⁴, na ktorom je založený aj test vierohodnostným pomerom:

$$LR = -2 \log \left(\frac{L_0}{L_f} \right) = 2(\log L_f - \log L_0)$$

²⁴ Veličinu $G^2 = -2 \log L$ Fox (2007) [15.] nazýva *reziduálna deviácia* a je zovšeobecnením reziduálnej sumy štvorcov v lineárnom regresnom modeli. Pozri tiež v Leblang 2004 [16.].

kde L_0 je vierohodnostná funkcia nulového modelu a L_f je vierohodnostná funkcia úplného modelu. Výsledná hodnota štatistiky pravdepodobnostného pomeru leží v intervale $(0, \infty)$. Čím je hodnota štatistiky väčšia, tým testovaný úplný model lepšie opisuje dané dáta.

2. V druhom prípade porovnáваме zvolený (úplný) model s modelom s určitými reštrikciami, ktorý má niektoré parametre rovné nule a predstavuje submodel úplného modelu. Hypotéza pre testovanie takýchto modelov je formulovaná nasledovne:

- H_0 : úplný model sa rovná submodelu, teda nie je štatisticky významnejší ako submodel a má niektoré parametre β_i rovné nule
- H_1 : úplný model sa nerovná submodelu a je štatisticky významnejší ako submodel

Zoberme si napríklad dva logistické modely. O jednom predpokladajme, že má q koeficientov rovných nule:

- $\text{logit}_f(\pi) = \beta_0 + \beta_1 X_1 + \dots + \beta_q X_q + \beta_{q+1} X_{q+1} + \dots + \beta_k X_k$
- $\text{logit}_r(\pi) = \beta_0 + 0X_1 + \dots + 0X_q + \beta_{q+1} X_{q+1} + \dots + \beta_k X_k = \beta_{q+1} X_{q+1} + \dots + \beta_k X_k$

kde $\text{logit}_f(\pi)$ je označenie pre logit úplného modelu a $\text{logit}_r(\pi)$ predstavuje logit modelu s reštrikciami.

Pre oba modely vypočítame maximalizovanú vierohodnostnú funkciu: L_f pre úplný model a L_r pre model s reštrikciami. Keďže model s reštrikciou je submodelom úplného modelu, platí:

$$L_f \geq L_r$$

Ak na modely aplikujeme test vierohodnostným pomerom, dostaneme LR štatistiku, ktorá má χ_q^2 rozdelenie, kde q predstavuje počet stupňov voľnosti, rovnajúci sa počtu nulových parametrov:

$$LR = -2 \log \left(\frac{L_r}{L_f} \right) = 2(\log L_f - \log L_r) \sim \chi_q^2$$

Nulovú hypotézu H_0 zamietame, ak má logaritmická štatistika testu vierohodnostným pomerom na zvolenej hladine významnosti väčšiu hodnotu ako kvantil χ_q^2 rozdelenia.

Test vierohodnostným pomerom môžeme použiť aj v prípade zisťovania, či pridanie ďalšej nezávislej premennej nezlepší vysvetlenie závislej premennej.

1.4.9 Miera priliehavosti v Logit a Probit modeloch

Nakoľko sa parametre v regresných modeloch s binárnou závislou premennou odhadujú pomocou maximalizácie vierohodnostnej funkcie, na určenie miery priliehavosti modelu nemôžeme použiť „klasický“ koeficient determinácie R^2 z lineárneho regresného modelu²⁵. Za účelom vypočítania, do akej miery je skúmaný model vhodný pre popis daných dát, však boli vyvinuté viaceré analogické štatistiky, nazývané *pseudo* R^2 ²⁶, s podobnými vlastnosťami ako koeficient determinácie R^2 . Hodnoty týchto štatistík sa väčšinou pohybujú v rozmedzí intervalu (0,1) a v platnosti zostáva aj tvrdenie, že čím je hodnota štatistiky vyššia, tým odhadovaný model lepšie popisuje dané data.

McFadden-ova štatistika *pseudo* R^2 ²⁷ využíva hodnotu logaritmu vierohodnostnej funkcie úplného modelu l_f a hodnotu logaritmu vierohodnostnej funkcie modelu s konštantným parametrom l_0 :

$$R_{pseudo}^2 = 1 - \frac{\log L_f}{\log L_0} = 1 - \frac{l_f}{l_0}$$

Čím je logaritmus vierohodnostnej funkcie l_f modelu s parametrami menší v porovnaní s l_0 modelu s konštantou, tým je pomer logaritmov vierohodnostných funkcií bližšie k 0, miera priliehavosti je vyššia a úplný model lepšie vystihuje modelované dáta.

²⁵ Koeficient determinácie je popísaný aj v Hatrák 2007 [2.].

²⁶ Murphy 2007 [17.].

²⁷ Podľa Croux 2007 [14.], Siebertová 2005 [18.], EViews 2004 [21.].

1.4.10 Miera správnosti predikcie

Ak máme odhadnuté hodnoty všetkých parametrov modelu, vypočítame odhady závislej premennej, resp. pravdepodobnosť π_i pre všetky pozorovania dosadením do rovnice pre jednotlivé modely:

$$\pi_i = F(\hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_k X_{ik})$$

Ak bude výsledná pravdepodobnosť väčšia ako 50%, predpovedáme $\hat{y}_i = 1$, a naopak, ak bude výsledná pravdepodobnosť menšia ako 50%, predpovedáme $\hat{y}_i = 0$. Výsledky porovnáme s pozorovanými dátami premennej y a vypočítame percentuálnu úspešnosť predikcie.

2 Trh práce na Slovensku

2.1 Historický vývoj

Vývoj na trhu práce odzrkadľuje realizáciu hospodárskej a sociálnej politiky štátu. Slovenský trh práce v rámci svojho historického vývoja prešiel rôznymi etapami - od samotného vzniku samostatného trhu práce a obdobia transformácie²⁸ slovenskej ekonomiky v 90. tých rokoch, cez obdobie prudkého nárastu nezamestnanosti, až po súčasný progresívny stav.

Výrazná zmena na trhu práce sa udiala v roku 1998 po zmene smerovania hospodárskej politiky štátu, ktorá si dala za cieľ zastavenie nerovnovážneho ekonomického rastu²⁹ z predchádzajúcich troch rokov (1996-1998). Dôsledok nového smerovania hospodárskej politiky sa prejavil už v roku 1999 v náraste miery nezamestnanosti o 3,7%³⁰ (najväčší medziročný nárast od roku 1993 predstavoval skoro stotisíc nezamestnaných) na hodnotu 16,2% (416 800 osôb). Nezanedbateľný podiel na tomto nepriaznivom stave malo zameranie sa na rast produktivity práce, reštriktívne opatrenia vo verejnom rozpočte, sanácia vysokého deficitu obchodnej bilancie, systémová reštrukturalizácia strategických podnikov atď., čo sa odzrkadlilo vo veľkom prepúšťaní, najmä pracovníkov s nižším vzdelaním a vekovo starších. V roku 2001 dosiahla úroveň miery nezamestnanosti na Slovensku svoje historické maximum 19,2%, čo predstavovalo vyše pol milióna ľudí bez práce (pozri graf 2.1 a 2.2). Takýto negatívny trend sa udržal aj nasledujúce roky. Až od roku 2004 má nezamestnanosť na Slovensku trvalo klesajúci charakter. V roku 2005 bola zaznamenaná hodnota miery nezamestnanosti na úrovni 16,2%, čo bol medziročný pokles o 1,9%. Podmienky pre tento zlepšený výsledok sa vytvorili po vstupe Slovenska do Európskej únie. Príliv zahraničných investícií, najmä do automobilového priemyslu a s tým spojený synergický efekt dodávateľov, subdodávateľov a tvorby infraštruktúry, znamenal tvorbu nových pracovných miest. V rokoch 2006 a 2007 pokračovala klesajúca tendencia miery nezamestnanosti, ktorá dosiahla hodnoty 13,3%, resp. 11,0%, čo je doposiaľ najnižšia hodnota v histórii samostatnosti Slovenska a predstavuje niečo vyše 290 tisíc osôb bez práce (pozri graf 2.1 a 2.2).

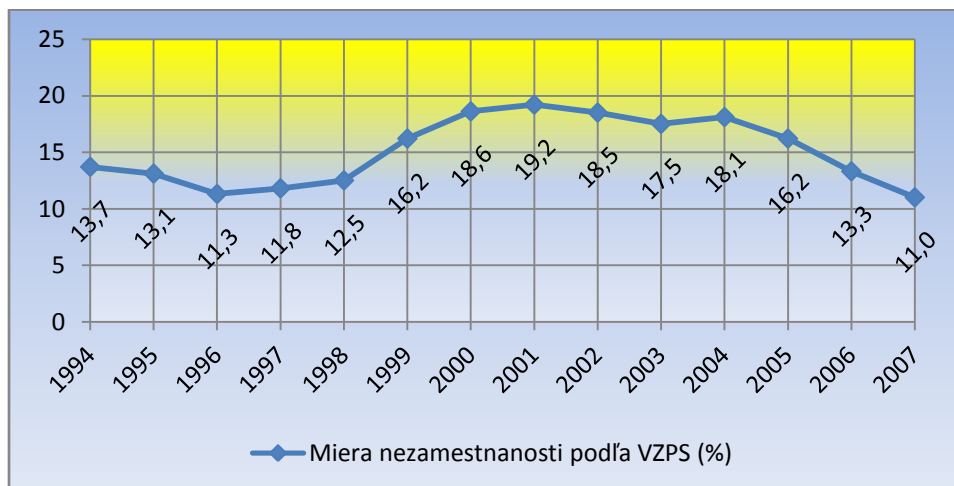
²⁸ Myslí sa transformácia z centrálne plánovanej na trhovú ekonomiku. Kvetan V., Páleník V., a spoluautori 2005 [7.]

²⁹ Ako uvádza Kvetan V., Páleník V., a spoluautori 2005 [7.].

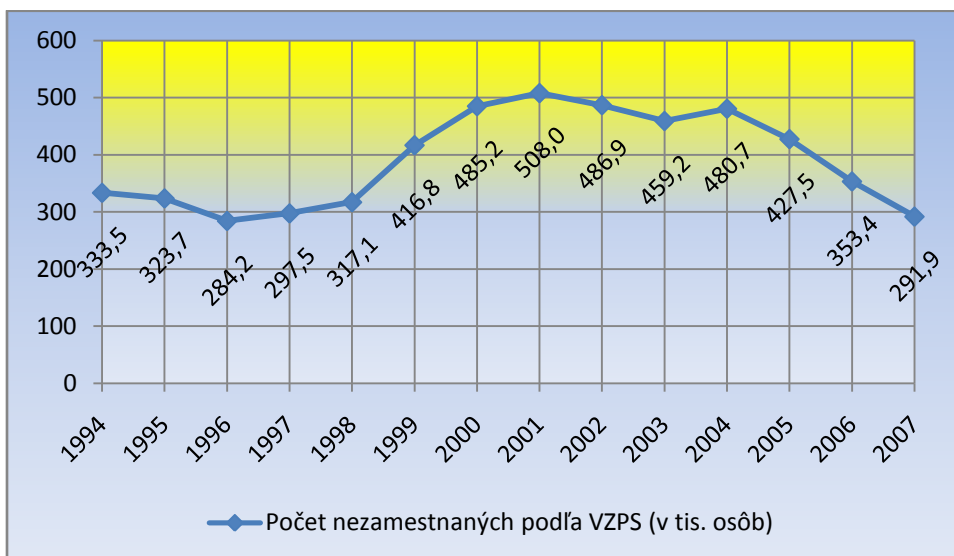
³⁰ Údaje podľa zistení ŠÚ SR, metódou VZPS.

Hlavným dôvodom pokračujúceho znižovania miery nezamestnanosti aj naďalej zostáva prílev ďalších nových zahraničných investícií, s tým súvisiaci synergický efekt, celkové zlepšenie ekonomického prostredia, ale aj odchod mladých ľudí za lepšie platenou prácou do zahraničia³¹. Otázkou zostáva, ako rýchlo a či vôbec je Slovensko schopné dosiahnuť prirodzenú mieru nezamestnanosti³² vzhľadom na niektoré pretrvávajúce problémy trhu práce.

Graf 2.1: Vývoj miery nezamestnanosti v SR³³



Graf 2.2: Vývoj počtu nezamestnaných v SR



³¹ Vo Výberovom zisťovaní pracovných síl sú ľudia vykonávajúci prácu v zahraničí, patriaci do zisťovanej domácnosti, evidovaní v zastúpení iného člena domácnosti a započítavaní medzi pracujúcich

³² Prirodzená miera nezamestnanosti (NAIRU) je pre Slovensko vypočítaná na úrovni 6% (Morvay 2005).

³³ Dáta pre grafy týkajúce sa trhu práce pochádzajú z verejnej bázy dát <http://www.statistics.sk/pls/elisw/vbd> a Ekonomického ústavu Slovenskej akadémie vied.

2.2 Problémy trhu práce na Slovensku³⁴

Trh práce môžeme definovať ako určitý priestor, na ktorom sa stretáva ponuka a dopyt po ľudskej pracovnej sile a na ktorom môže dochádzať v dôsledku jeho otvorenosti k vzniku určitej nevyváženosti vo vzťahu ponuky a dopytu.

Od začiatku 90. rokov možno na Slovensku pozorovať stabilné črty previsu ponuky pracovnej sily nad dopytom, od čoho sa následne odvíja sociálne rozvrstvenie spoločnosti, jej príjmová štruktúra a využívanie existujúceho ľudského kapitálu.

Hlavným a pretrvávajúcim problémom trhu práce na Slovensku je vysoká miera nezamestnanosti, veľká skupina dlhodobo nezamestnateľnej časti obyvateľstva i neuspokojený dopyt po výkonnej a kvalifikovanej pracovnej sile. V poslednom čase sa výraznejším problémom stala nezamestnanosť mladých ľudí a narušené väzby medzi vzdelanostnou úrovňou a nárokmi trhu práce.

Negatívnym javom objavujúcim sa v časti populácie na Slovensku je klesajúca snaha o aktívne uplatňovanie sa na trhu práce. Na strane ponuky práce stále chýba dostatočná ochota k zvyšovaniu vzdelania, zmene kvalifikácie, adaptabilita a mobilita pracovnej sily. Problémom po príchode zahraničných investorov sa ukazuje dlhší reakčný čas školského systému na zmenené požiadavky trhu práce. Prekvapujúco nízka je tiež zamestnanosť mužskej populácie v produktívnom veku. (Azudová 2005)

Výrazným a pretrvávajúcim problémom je nerovnováha na regionálnych trhoch práce, ktorá je daná historickým, kultúrnym, geografickým a ekonomickým vývojom. Nízka medziregionálna mobilita pracovnej sily vytvára kvalitatívne neproporcionálne rozvrstvenie pracovnej sily v rôznych častiach krajiny. Väčšina ekonomických aktivít sa sústreďuje do hlavných mestských oblastí a priemyselných centier so stabilizovanou mierou ponuky a dopytu po pracovnej sile. Na druhej strane však vznikajú oblasti s vysokou nezamestnanosťou a absenciou pracovných príležitostí. Veľké aglomerácie, v ktorých sa vo väčšej miere koncentruje vysokokvalifikovaná pracovná sila, vykazujú podstatne vyššiu zamestnanosť a vyššie príjmy než ostatné oblasti. V poslednom období však pretrváva snaha vlády o postupné vyrovňovanie regionálnej diferencovanosti územia Slovenska.

³⁴ Inšpirované Azudová 2005 [1.].

Od začiatku tohto storočia sa prejavujú obavy zo zvyšovania počtu nezamestnaných v sociálne rizikovej skupine mladých ľudí. V roku 2001 tvorila skupina mladých ľudí medzi 15-29 rokov (asi 230 tisíc osôb) 45,4% z celkového počtu nezamestnaných, čo bol dovtedy najvyšší pomer v histórii slovenského trhu práce (pozri graf 2.3 a 2.4). Postupom času sa vplyvom zvyšovania tvorby pracovných miest, technologického rozvoja a hrozby medzigeneračného konfliktu situácia zlepšovala a v roku 2007 tvorila skupina mladých ľudí medzi 15-29 rokov (asi 102 tisíc osôb) už len 34,8% z celkového počtu nezamestnaných.

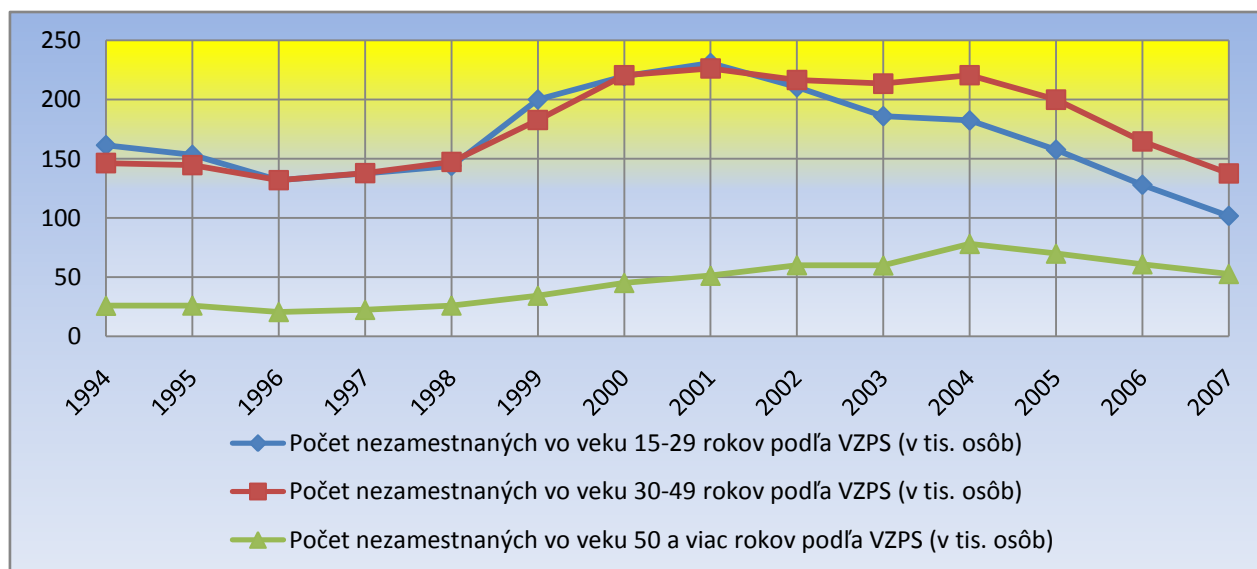
Skúsenosti z vývoja trhu práce v ostatnom období naznačujú istý trend, keď mladí ľudia dávajú prednosť slobodnému pohybu medzi nezamestnanosťou, dočasnou prácou, krátkodobou migráciou do zahraničia a pod. pred voľbou usadenia sa a prijatia nižšie platenej a menej kvalifikovanej práce, ako by si predstavovali. Uspokojivé sociálne zázemie v niektorých rodinách vedie tiež časť mladých nezamestnaných k odkladu vstupu na trh práce dovtedy, kým im to bude osobne vyhovovať, resp. kým to bude nevyhnutne potrebné.

Počet nezamestnaných ľudí vo veku 50 rokov a starších od roku 1994 v podstate neustále rástol (pozri graf 2.3 a 2.4). V roku 2004 dosiahol svoje maximum na hodnote vyše 78 tisíc nezamestnaných, čo bol trojnásobný nárast oproti spomínanému roku 1994. Po roku 2004 sa začala situácia zlepšovať, avšak pomer starších nezamestnaných ľudí k celkovému počtu nezamestnaných ďalej stúpala, čo svedčí o nižšom tempe poklesu počtu nezamestnaných v tejto vekovej kategórii oproti ostatným skupinám.

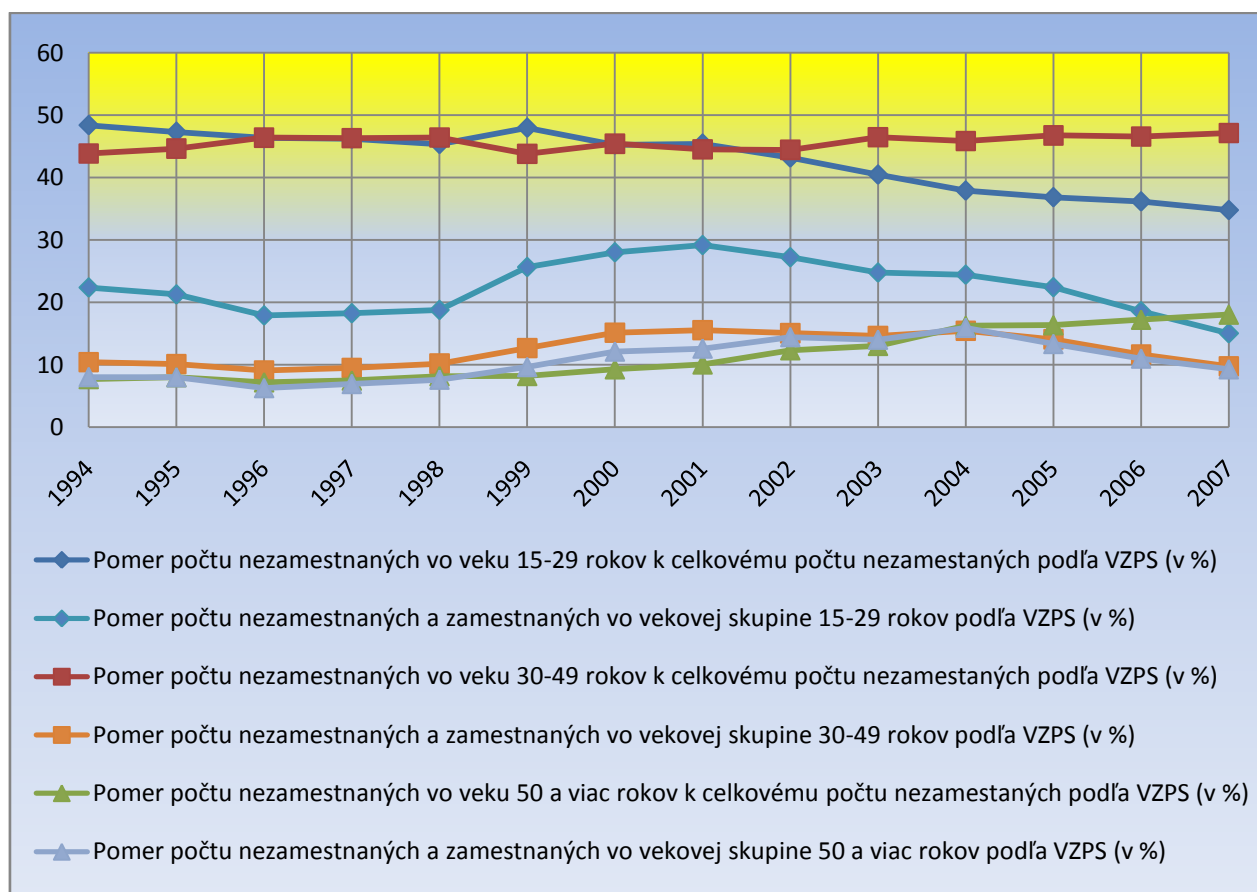
Azudová (2005) uvádza, že miera vzdelania vykazuje z pohľadu úspešného uplatnenia sa na trhu práce koreláciu s požiadavkami trhu práce. Tiež tvrdí, že v skupine vysokoškolsky vzdelaných ľudí si pomerne viac ľudí nájde prácu, ako v skupine s nižším vzdelaním.

Pri porovnaní (pozri graf 2.5 a 2.6) roku 1996, kedy bola zaznamenaná najnižšia miera nezamestnanosti v 90. tých rokoch (11,3 %, čo predstavovalo asi 284 tisíc nezamestnaných osôb), roku 2001, kedy bola najvyššia miera nezamestnanosti (19,2% a počet nezamestnaných narástol na 508 tisíc osôb) a roku 2007 (11,0% - 292 tisíc osôb), kedy sa skoro všetky zložky miery nezamestnanosti dostali približne na úroveň roku 1996, si môžeme všimnúť zaujímavý vývoj v skupine so základným a vysokým vzdelaním. Postupne, počas týchto rokov, sa počet ľudí so základným vzdelaním výrazne znižoval, ale rástol pomer nezamestnaných so základným vzdelaním. Počet ľudí s vysokým vzdelaním rástol, ale pomer nezamestnaných v tejto kategórii len kolísal.

Graf 2.3: Vývoj počtu nezamestnaných v SR podľa vekových skupín



Graf 2.4: Pomer počtu nezamestnaných v SR podľa vekových skupín



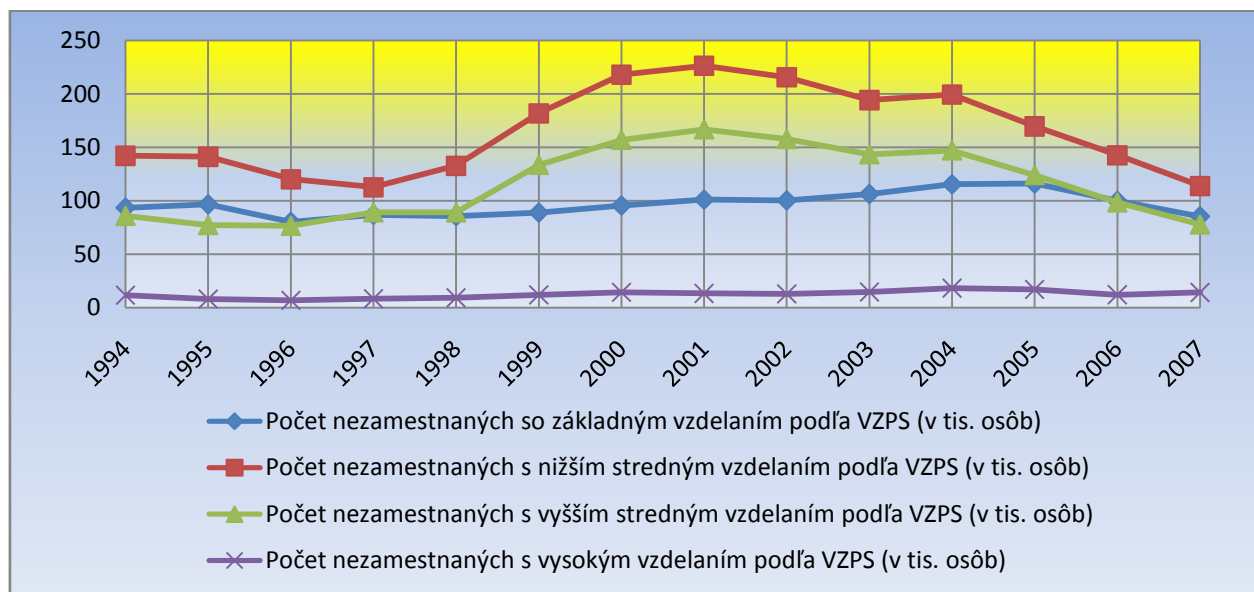
V roku 1999 bol zaznamenaný počet nezamestnaných so základným vzdelaním okolo 80 tisíc osôb (28,3% z celkového počtu nezamestnaných a 25,2% z celkového počtu základne vzdelaných). Tak ako aj v iných vzdelanostných skupinách stúpol počet nezamestnaných v roku 2001 aj v skupine základne vzdelaných na hodnotu asi 101 tisíc osôb (19,9% z celkového počtu nezamestnaných a 42,9% z celkového počtu základne vzdelaných). V roku 2007 bol zaznamenaný počet nezamestnaných so základným vzdelaním na úrovni asi 85,5 tisíc osôb (skoro 29,3% z celkového počtu nezamestnaných a 44,7% z celkového počtu základne vzdelaných).

V roku 1999 bol zaznamenaný počet nezamestnaných s vysokoškolským vzdelaním okolo 7 tisíc osôb (2,5% z celkového počtu nezamestnaných a 2,7% z celkového počtu vysokoškolsky vzdelaných). Počet nezamestnaných v roku 2001 v skupine vysokoškolsky vzdelaných stúpol skoro dvojnásobne, na hodnotu asi 13 400 osôb (2,6% z celkového počtu nezamestnaných a 4,8% z celkového počtu vysokoškolsky vzdelaných). V roku 2007 bol zaznamenaný počet nezamestnaných s vysokoškolským vzdelaním na úrovni asi 14 375 osôb (skoro 5% z celkového počtu nezamestnaných a 3,8% z celkového počtu vysokoškolsky vzdelaných).

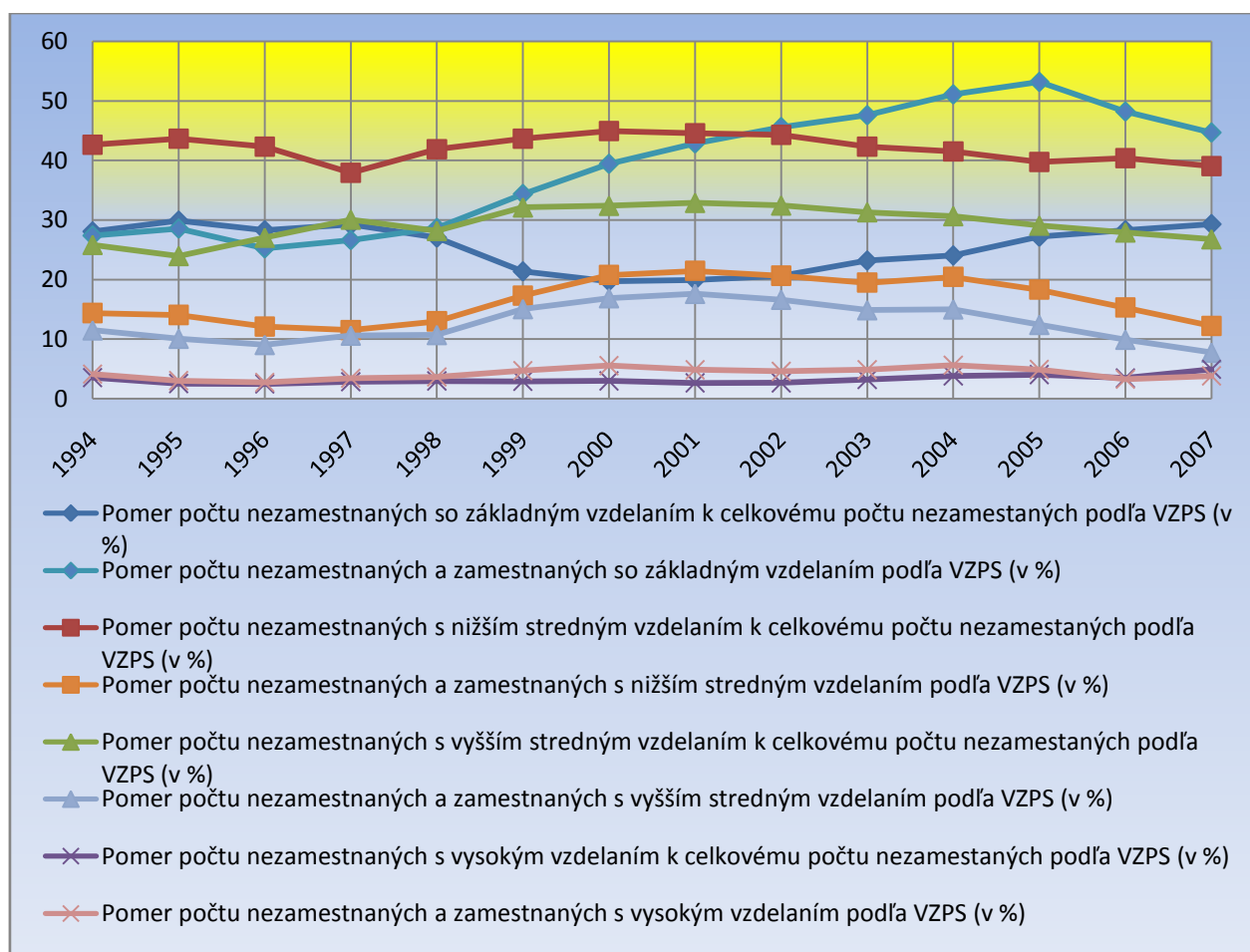
Pri realizácii aktívnej politiky trhu práce treba vytvárať a prijímať také opatrenia, ktoré by napomáhali mladej a menej vzdelanej pracovnej sile vyrovnávať šance na umiestnenie sa na oficiálnom trhu práce, aby týmto dvom najrizikovejším kategóriám nehrozilo nebezpečenstvo dlhodobej nezamestnanosti a vylúčenia z trhu práce. Skúsenosti z prvého zamestnania majú totiž nenahraditeľný socializačno-formatívny význam.

Pozitívne trendy, ktoré sa postupne prejavili v poklese nezamestnanosti po roku 2001 a výrazne hlavne za posledné dva roky, sa až natoľko neodrazili na dlhodobej nezamestnanosti. Veľká skupina dlhodobo nezamestnateľnej časti obyvateľstva, ktorá sa skladá prevažne z nízkokvalifikovanej pracovnej sily, je dôsledkom regionálnej diferencovanosti, nízkej vzdelanostnej štruktúry, nízkej mobility a flexibility určitej časti obyvateľstva. Jedným z ďalších dôvodov je určite aj dobrovoľné prijatie statusu nezamestnaného a snaha zotrvať v tejto pozícii, čo bolo v minulosti dlhodobo podporované poskytovaním podpory v nezamestnanosti, ktorá bola z pohľadu hľadania práce demotivujúcim faktorom.

Graf 2.5: Vývoj počtu nezamestnaných v SR podľa vzdelania

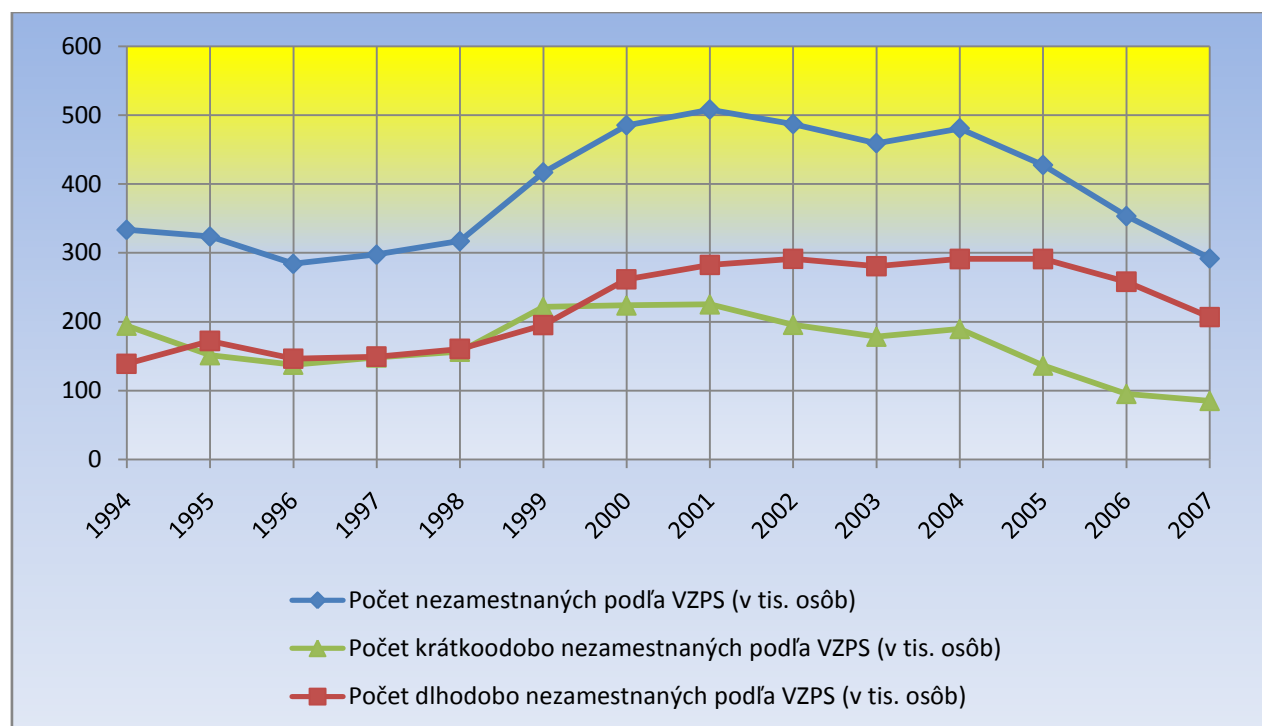


Graf 2.6: Pomer počtu nezamestnaných v SR podľa vzdelania

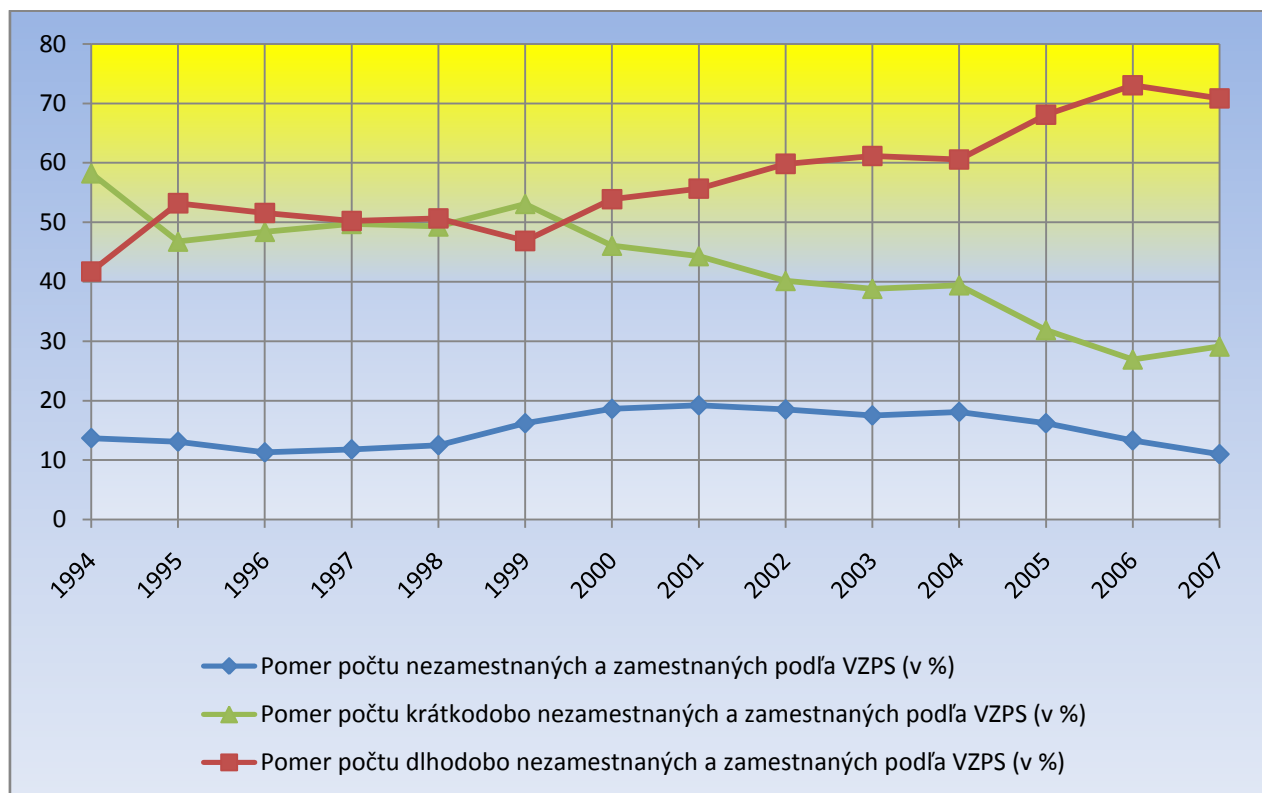


Počet dlhodobo nezamestnaných (viac ako 1 rok) v roku 2002 dosahoval vyše 291 tisíc, čo bola najvyššia hodnota od roku 1994 a pomerne to predstavovalo 59,8% z celkového počtu nezamestnaných (pozri graf 2.7 a 2.8). V ďalších rokoch sa počet dlhodobo nezamestnaných a ich percentuálny podiel udržiaval približne na rovnakej úrovni, čo svedčí o určitých stabilných vlastnostiach trhu práce, neschopnosti absorbovať časť existujúcej štruktúry pracovnej sily a zároveň o prejavujúcej sa pasivite a dobrovoľnom opúšťaní ambícií zamestnať sa. Až v roku 2006 počet dlhodobo nezamestnaných výraznejšie klesol. Zároveň sa však znížil aj celkový počet nezamestnaných a tak podiel dlhodobo nezamestnaných dosiahol rekordne vysokú hodnotu 73,1% (porovnanie s rokom 1994 - 41,7%). V roku 2007 klesol opäť celkový počet nezamestnaných, čo sa prejavilo aj na dlhodobej nezamestnanosti. Avšak pomer týchto dvoch ukazovateľov zostal stále vysoký (70,9%).

Graf 2.7: Vývoj počtu nezamestnaných v SR podľa dĺžky nezamestnanosti



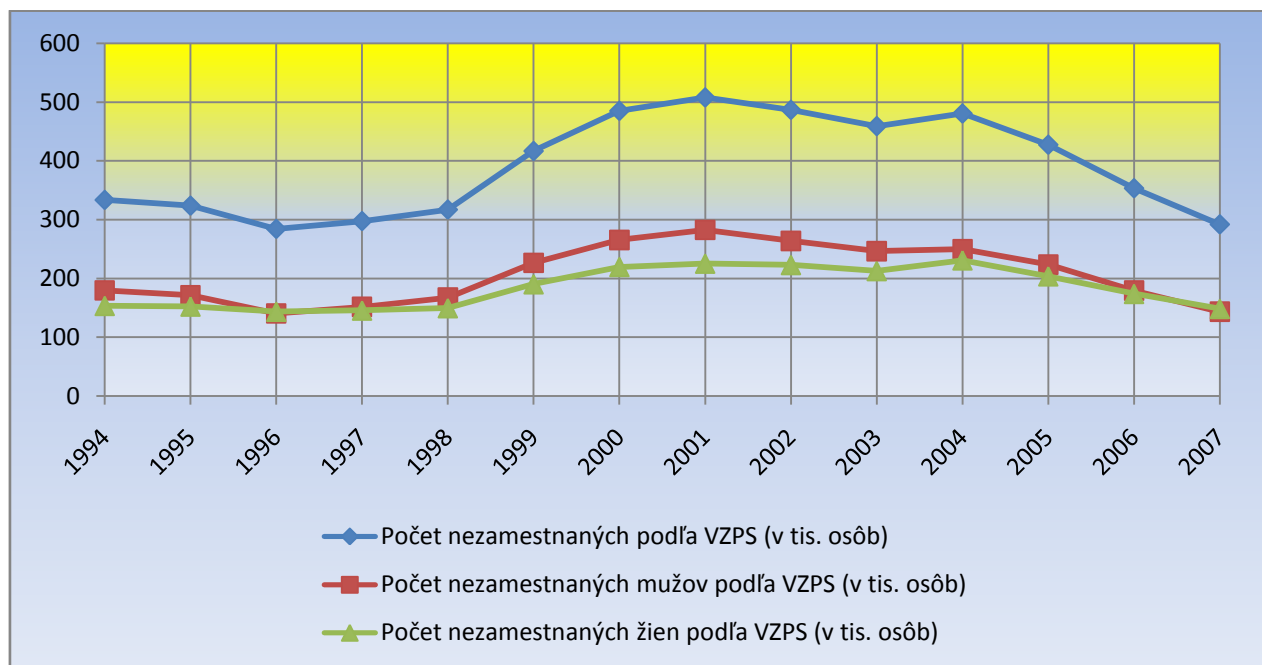
Graf 2.8: Pomer nezamestnaných v SR podľa dĺžky nezamestnanosti



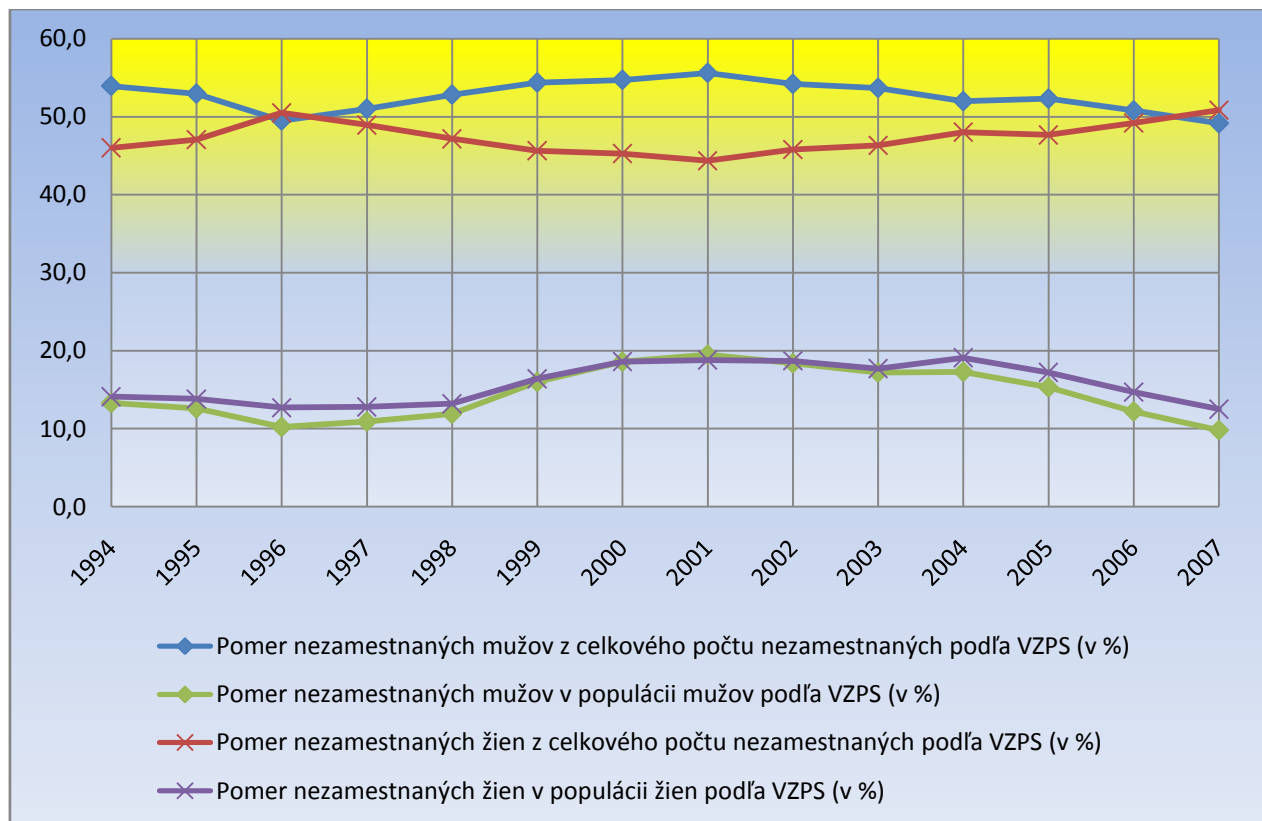
V demokratickej spoločnosti by mal mať uchádzač o pracovnú pozíciu rovnakú šancu získať prácu bez ohľadu na rodovú príslušnosť. V praxi sa ale môže často stať, že zamestnávateľia uprednostnia mužského kandidáta. Tento skrytý diskriminačný faktor sa môže prejavovať aj pri pohľade na nezamestnanosť podľa pohlavia (pozri graf 2.9 a 2.10). Z grafu však jasne vyplýva, že za posledných desať rokov bol vyšší počet nezamestnaných mužov ako žien, ale pomer nezamestnaných mužov a žien vo svojej kategórii bol približne vyrovnaný. V posledných troch rokoch môžeme vidieť v skupine žien nižšiu tendenciu poklesu miery nezamestnanosti ako v skupine mužov. V roku 2007 dokonca pomer počtu nezamestnaných žien k počtu nezamestnaných mužov presiahol 50 %-nú hranicu.

Ako sme mohli postupne sledovať, štruktúra trhu práce na Slovensku závisí od viacerých činiteľov a vývoj trhu práce nie je úplne stabilný. Každý vážnejší zásah do tohto procesu môže preto ovplyvniť výšku miery nezamestnanosti a štrukturálne zloženie skupiny nezamestnaných na niekoľko rokov do budúcnosti.

Graf 2.9: Vývoj počtu nezamestnaných mužov a žien v SR



Graf 2.10: Pomer nezamestnaných mužov a žien v SR



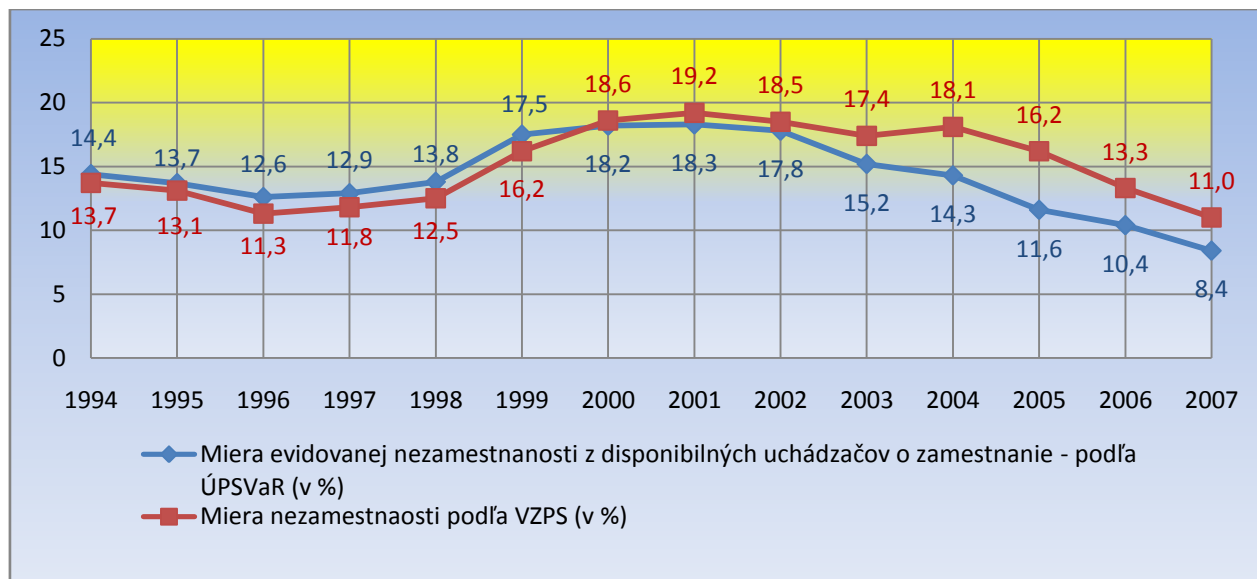
2.3 Zber informácií o trhu práce na Slovensku

2.3.1 Metodiky zberu informácií o trhu práce

V súčasnosti sa na Slovensku používajú dva rozdielne prístupy k zberu informácií o trhu práce: Výberové zisťovanie pracovných síl - metodika Štatistického úradu SR a štatistiky vykazované Ústredím práce, sociálnych vecí a rodiny.

Výsledky Výberových zisťovaní pracovných síl slúžia najmä pre potreby medzinárodnej štatistiky práce a na medzinárodné porovnania. Umožňujú tiež objektívnejší pohľad na stav a vývoj nezamestnanosti z hľadiska vývoja vekovej a kvalifikačnej štruktúry nezamestnaných, dĺžky trvania nezamestnanosti, štruktúry podľa odvetví a zamestnaní, a sú východiskom analýz i prognóz ekonomického a sociálneho rozvoja v SR. Údaje úradov práce sú významným indikátorom pre uplatňovanie politiky trhu práce. Oba prístupy majú svoje prednosti i nedostatky. Údaje získané oboma spôsobmi môžeme ponímať ako dva navzájom sa dopĺňajúce zdroje informácií o stave a vývoji trhu práce v SR. (Kotulič 2007) (pozri graf 2.11)

Graf 2.11: Porovnanie vývoja miery nezamestnanosti v SR



Ďalej sa budeme zaoberať len metodikou VZPS, podľa ktorej sa osoby zaraďujú medzi pracujúcich a nezamestnaných v súlade s medzinárodnými definíciami a odporúčaniami Medzinárodnej organizácie práce (ILO).

2.3.2 Výberové zisťovanie pracovných síl (VZPS) ³⁵

2.3.2.1 Popis metodiky VZPS

Výberové zisťovanie pracovných síl je metóda zberu informácií o stave, štruktúre a vývoji trhu práce. Štatistický úrad SR ju využíva pri meraní miery zamestnanosti, nezamestnanosti a ekonomickej aktivity obyvateľstva na Slovensku. VZPS je štandardnou metódou odporúčanou medzinárodnými inštitúciami pod patronátom Medzinárodnej organizácie práce. Na našom území prebieha zisťovanie kontinuálne od roku 1993, hoci spočiatku pod iným názvom a s pozmenenou metodikou. Štatistický úrad spracováva a vydáva publikácie obsahujúce tabuľky so štvrtročnými a ročnými výsledkami VZPS.

Základom pre výberové zisťovanie pracovných síl je náhodný výber pozostávajúci z 10250 bytov, ktoré predstavujú približne 0,6% z trvalo obývaných domácností v SR. Výberový súbor sa zostavuje štvrtročne a pokrýva všetky okresy SR. V súčasnosti sa používa proporcionálna metodika výberu bytov, podľa ktorej je rozsah výberového súboru závislý od veľkosti okresu danej počtom jeho obyvateľov. V tejto metodike sa zohľadňuje územno-správne usporiadanie Slovenska z roku 1996. Výber teda prebieha nezávisle v každom okrese.

Výberový súbor je pravidelne obmieňaný. Každý byt je súčasťou zisťovania päť za sebou idúcich štvrtrokov. Štvrtročne sa obmieňa 20% vzorky. Súbor bytov pre daný štvrtrok zisťovania je rovnomerne rozložený do jednotlivých týždňov toho štvrtroka.

Vlastný proces zisťovania informácií prebieha prostredníctvom osobných návštev alebo telefonického kontaktu pracovníkov Štatistického úradu SR s domácnosťami zastúpenými vo vzorke. Miera úspešnosti zberu dát sa pohybuje na úrovni okolo 94%.

Predmetom výberového zisťovania pracovných síl sú osoby vo veku od 15 rokov žijúce v daných domácnostiach, bez ohľadu na to, či majú v byte trvalý, prechodný alebo nehlásený (dočasný) pobyt, okrem inštitucionálneho obyvateľstva (t.j. osôb vo výkone trestu, nesvojprávnych osôb umiestnených v ústavoch a pod.).

Údaje získané od respondentov sú na základe matematicko-štatistických metód prepočítané na demografické údaje o stave obyvateľstva Slovenskej republiky k 1.1. sledovaného

³⁵ Podľa publikácie ŠÚ SR - *Výberové zisťovanie pracovných síl* (1994-2001) [13.].

kalendárneho roka podľa územnej jednotky, pohlavia a vekovej skupiny. Výsledné štandardné chyby odhadov skutočných ukazovateľov pre celú populáciu predstavujú v prípade zamestnaných 0,3% - 0,6%, u nezamestnaných 1% - 2,5% a pri miere nezamestnanosti 0,2% - 0,4%.

2.3.2.2 Metodické vymedzenie ukazovateľov³⁶

Pracujúci – za pracujúcich sú považované všetky osoby vo veku od 15 rokov, ktoré v referenčnom týždni vykonávajú aspoň jednu hodinu prácu za mzdu, plat alebo prácu za účelom dosiahnutia zisku, vrátane osôb pracujúcich v zahraničí. Môže ísť o prácu na plný alebo skrátený úväzok, dočasnú, príležitostnú alebo sezónnu prácu. Medzi pracujúcich sú zahrnutí aj vypomáhajúci členovia domácností podnikateľov, ktorí za svoju činnosť nepoberajú žiadnu mzdu ani odmenu, profesionálni príslušníci ozbrojených zložiek, ako aj osoby v civilnej službe. Za pracujúcich sa považujú tiež osoby, ktoré majú prácu, ale v sledovanom týždni nepracujú pre chorobu, dovolenku, riadnu materskú dovolenku, školenie, zlé počasie, v dôsledku štrajku a výluky, s výnimkou osôb na dlhodobom neplatenom voľne a osôb na ďalšej materskej dovolenke.

Nezamestnaní - za nezamestnaných sú považované všetky osoby vo veku od 15 rokov, ktoré súčasne spĺňajú tri podmienky:

1. v sledovanom týždni nemajú žiadnu platenú prácu
2. v posledných štyroch týždňoch si aktívne hľadajú prácu
3. sú schopné nastúpiť do práce najneskôr do dvoch týždňov

Tieto osoby môžu, ale nemusia byť evidované na úradoch práce ako nezamestnané.

Za nezamestnané sa považujú aj nepracujúce osoby, ktoré si prácu našli a do zamestnania nastúpia v priebehu jedného mesiaca.

Od roku 1997 sú medzi nezamestnaných zaradené aj osoby na dlhodobom neplatenom voľne, ktoré si v posledných štyroch týždňoch aktívne hľadajú prácu a sú schopné nastúpiť do práce najneskôr do dvoch týždňov, alebo ktoré si prácu nehľadajú z toho dôvodu, že v blízkej budúcnosti očakávajú návrat do svojho zamestnania.

³⁶ Definície podľa Medzinárodnej organizácie práce (ILO).

3 Modelovanie úspešnosti na trhu práce

3.1 Príprava modelovania

3.1.1 Výber individuálnych dát

Cieľom tejto diplomovej práce bolo modelovanie úspešnosti pri hľadaní zamestnania, pri ktorom sme potrebovali použiť dáta pochádzajúce z trhu práce. Vyvinuli sme snahu získať potrebné individuálne dáta z Výberového zisťovania pracovných síl priamo zo Štatistického úradu Slovenskej republiky. Požadované dáta sa však nepodarilo získať kvôli neochote a finančným požiadavkám Štatistického úradu. Na modelovanie boli nakoniec použité upravené individuálne dáta získané na základe VZPS, poskytnuté Inštitútom zamestnanosti na účel vypracovania tejto diplomovej práce.

Našou prvotnou snahou pri skúmaní úspešnosti na trhu práce bolo porovnať stav trhu práce v dvoch dostatočne vzdialených časových obdobiach. Preto sme sa na jednej strane snažili získať čo najaktuálnejšie individuálne dáta pochádzajúce z VZPS, najlepšie za rok 2007. Chceli sme ich porovnať s analogickými dátami za obdobie približne pred desiatich rokov. Počas písania diplomovej práce však ešte neboli k dispozícii kompletne individuálne dáta za rok 2007, a tak sme si zvolili predchádzajúci rok 2006. Na druhej strane sme pre porovnanie potrebovali získať aj staršie dáta. Kvôli zmene metodiky vo VZPS v roku 1999 a ďalším problémom, ktoré sa vyskytli pri agregovaní dát na Inštitúte zamestnanosti, sa nám však nepodarilo získať potrebné dáta za obdobie pred desiatich rokov. Ako druhé obdobie pre porovnanie sme si preto vybrali rok 2000. V tomto roku sa začali naplno prejavovať dôsledky reštriktívnych opatrení vtedajšej vlády a nezamestnanosť prudko stúpala.

Nakoniec sme teda pracovali so štvrťročnými dátami za roky 2000 a 2006. Dáta za jednotlivé kalendárne štvrťroky daných rokov sme zaradili do jednej vzorky a využívali sme ich ako rôzne individuálne údaje za daný rok. Pri modelovaní stavu trhu práce za rok 2000 sme tak mali k dispozícii 59691 pozorovaní a za rok 2006 to bolo 53702 pozorovaní. Realizácia výpočtov prebiehala v prostredí štatistického softvéru EViews 6.0 a MS Excel 2007.

3.1.2 Voľba premenných modelu a formulácia hypotéz

Úspešnosť na trhu práce sme pôvodne chceli vyjadriť ako tranzíciu, teda prechod od statusu nezamestnaného k statusu zamestnaného, počas dvoch následných období a v závislosti od určitých parametrov trhu práce. Z dát, ktoré sme mali k dispozícii sa nám však nepodarilo získať informácie o tranzícii. Bolo to najmä z dôvodu nemožnosti identifikovať jednotlivca v dvoch po sebe nasledujúcich obdobiach. Ako vysvetľovanú premennú sme si zvolili štatút na trhu práce podľa definície ILO, teda či je respondent zamestnaný (pracujúci) alebo nezamestnaný. Vysvetľujúce premenné sme si zvolili na základe popisu vývoja a aktuálneho stavu na trhu práce. Zahrnuli sme medzi ne pohlavie, vekové skupiny, najvyššie dosiahnuté vzdelanie, miesto trvalého bydliska a rodinný stav.

Predpokladáme, že vek je dôležitým faktorom ovplyvňujúcim status osoby na trhu práce. Mladí ľudia v dnešnej dobe nemajú toľko záväzkov, chcú spoznávať svet a skúmať nové obzory. Trvalo sa zamestnávajú, až keď dostanú zaujímavú pracovnú ponuku alebo to budú osobne potrebovať. Starší ľudia sú menej flexibilní a ak stratia dlhoročnú prácu, majú menšiu šancu si nájsť adekvátnu náhradu. Z grafu 2.3 a 2.4 môžeme vidieť, že počet mladých nezamestnaných tvorí viac ako tretinu počtu všetkých nezamestnaných a počet starších nezamestnaných klesá len pomaly, čo sa prejavuje na zvyšovaní pomeru medzi všetkými nezamestnanými. Do modelu boli preto zaradené dve binárne premenné pre vekovú skupinu mladých (15-29 rokov) a starých (50 a viac rokov). Pri oboch premenných očakávame negatívny vplyv na status zamestnanosti. Teda ak človek patrí do mladšej alebo staršej vekovej kategórie, bude nižšia pravdepodobnosť, že je zamestnaný. Výsledky modelu potvrdia, prípadne vyvrátia predpoklady o rizikovosti týchto vekových skupín.

Ďalším dôležitým faktorom ovplyvňujúcim status osoby na trhu práce je výška dosiahnutého vzdelania, ktorá priamo súvisí s kvalitou ľudského kapitálu. Ak má človek vyššie vzdelanie, zvyšuje sa aj jeho všeobecný prehľad a má vyššiu šancu nájsť si prácu v širšom okruhu pracovných pozícií. Rovnako významným faktom je, že človek s vyšším dosiahnutým vzdelaním môže zaujať pracovné miesto, ktoré si vyžaduje nižšiu kvalifikáciu, kým opak je skôr výnimkou ako pravidlom. Z grafu 2.5 a 2.6 môžeme vidieť, že počet nezamestnaných so základným vzdelaním tvorí skoro 30% z celkového počtu nezamestnaných a zvyšuje sa. V skupine so základným vzdelaním je v súčasnosti takmer každý druhý človek nezamestnaný. Počet

nezamestnaných s vysokým vzdelaním naopak v poslednom období prekvapujúco stúpa. Do modelu sme sa preto rozhodli zahrnúť kvalitatívnu premennú zahŕňajúcu štyri vzdelanostné skupiny. Očakávame pozitívny vplyv tejto premennej. Čím má uchádzač o pracovnú pozíciu vyšší stupeň dosiahnutého vzdelania, tým je vyššia pravdepodobnosť, že bude zamestnaný.

Ďalšou binárnou premennou zahrnutou do modelu bolo pohlavie. Na Slovensku sa ešte stále v plnej miere neprejavuje rovnosť pohlaví pri výbere kandidátov do zamestnania. Zamestnávateľ bude určite veľmi dobre zvažovať, či prijme do zamestnania mladú nízko vzdelanú ženu. Z grafu 2.9 a 2.10 môžeme vyčítať, že pomer nezamestnaných je u žien vyšší ako u mužov, a v poslednom roku je medzi nezamestnanými viac žien ako mužov. Predpovedáme negatívny vplyv tejto premennej. Ak je respondentom žena, bude nižšia pravdepodobnosť, že je zamestnaná.

Miesto trvalého bydliska môže byť vďaka nízkej mobilite pracovnej sily na Slovensku dôležitým faktorom ovplyvňujúcim úspešnosť pri hľadaní zamestnania. Pri trvalom bydlisku v mestských aglomeráciách má uchádzač o prácu širší výber a väčšiu ponuku pracovných miest, a teda aj väčšiu šancu na nájdenie zamestnania. Ak je miestom trvalého bydliska vidiecka oblasť, prípadné dochádzanie do potencionálneho zamestnania, vzdialeného niekoľko desiatok kilometrov, môže byť náročné. Miesto trvalého bydliska preto predstavuje ďalšiu binárnu premennú nášho modelu. Pri tejto veličine očakávame opäť negatívny vplyv na status na trhu práce. Ak je miestom trvalého bydliska respondenta vidiecka oblasť, bude nižšia pravdepodobnosť, že je zamestnaný.

Poslednou premennou zahrnutou do skúmaného modelu je rodinný stav. Slobodný človek, ktorý nemá rodinu, nie je do takej miery viazaný, ako človek žijúci v manželstve. Môže si dovoliť byť istý čas nezamestnaný. Podobne sa môže správať aj rozvedený človek. Veličinu rodinný stav sme sa kvôli lepšej interpretovateľnosti výsledkov rozhodli rozdeliť na tri samostatné binárne premenné. Negatívny vplyv odhadujeme pre premennú popisujúcu slobodný a rozvedený stav a naopak pozitívny vplyv očakávame pri premennej manželského stavu.

3.1.3 Popis kvalitatívnych premenných

- Štatút na trhu práce - ILO:
 - 0 nezamestnaný podľa definície ILO
 - 1 pracujúci podľa definície ILO
- Pohlavie - SEX:
 - 0 muž
 - 1 žena
- Vekové skupiny:
 - YOUNG:
 - 0 osoby vo veku 30 a viac rokov
 - 1 osoby vo veku 15-29 rokov
 - OLD:
 - 0 osoby vo veku 15-49 rokov
 - 1 osoby vo veku 50 a viac rokov
- Najvyššie dosiahnuté vzdelanie - EDU:
 - 0 bez vzdelania, základné a nezistené
 - 1 nižšie stredné vzdelanie (učňovské, stredné bez maturity)
 - 2 vyššie stredné vzdelanie (učňovské s maturitou, úplné stredné všeobecné, úplné stredné odborné, vyššie odborné)
 - 3 vysoké vzdelanie (vysokoškolské 1.,2., a 3. stupeň)
- Miesto trvalého bydliska Mesto/Vidiel - C_C:
 - 0 obyvateľ okresu krajského mesta
 - 1 obyvateľ iného okresu
- Rodinný stav:
 - SINGLE:
 - 0 rodinný stav iný ako slobodný/á
 - 1 slobodný/á
 - MARRIED:
 - 0 rodinný stav iný ako ženatý/vydatá
 - 1 ženatý/vydatá
 - DIVORCED:

- 0 rodinný stav iný ako rozvedený/á
- 1 rozvedený/á

3.2 Formulácia všeobecného modelu

Vysvetľovaná premenná *ILO* bude v konkrétnych Logit a Probit modeloch predstavovať pravdepodobnosť, či je skúmaný respondent zamestnaný alebo nie. Všeobecný model obsahuje sedem binárnych vysvetľujúcich premenných, jednu kategóriálnu a môžeme ho zapísať nasledovným spôsobom:

3.2.1 Všeobecný model

$$ILO = f(SEX, YOUNG, OLD, EDU, C_C, SINGLE, MARRIED, DIVORCED, \varepsilon)$$

3.2.2 Logit model

$$\begin{aligned} \pi &= \Lambda(\beta_0 + \beta_1 SEX + \beta_2 YOUNG + \beta_3 OLD + \beta_4 EDU + \\ &+ \beta_5 C_C + \beta_6 SINGLE + \beta_7 MARRIED + \beta_8 DIVORCED) = \\ &= \frac{1}{1 + e^{-(\beta_0 + \beta_1 SEX + \beta_2 YOUNG + \beta_3 OLD + \beta_4 EDU + \beta_5 C_C + \beta_6 SINGLE + \beta_7 MARRIED + \beta_8 DIVORCED)}} \end{aligned}$$

3.2.3 Probit model

$$\begin{aligned} \pi &= \Phi(\beta_0 + \beta_1 SEX + \beta_2 YOUNG + \beta_3 OLD + \beta_4 EDU + \\ &+ \beta_5 C_C + \beta_6 SINGLE + \beta_7 MARRIED + \beta_8 DIVORCED) = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_0 + \beta_1 SEX + \beta_2 YOUNG + \beta_3 OLD + \beta_4 EDU + \beta_5 C_C + \beta_6 SINGLE + \beta_7 MARRIED + \beta_8 DIVORCED} e^{\left\{-\frac{1}{2}Z^2\right\}} dZ \end{aligned}$$

4 Analýza výsledkov

4.1 Úvod

Analýza výsledkov modelovania úspešnosti na trhu práce sa skladá z viacerých častí. Začneme od odhadov parametrov Logit a Probit modelov za roky 2000 a 2006. Keďže však samotné odhady parametrov nemajú pre nás veľkú výpovednú hodnotu, budú nás skôr zaujímať smery vplyvu jednotlivých odhadnutých parametrov, ktoré porovnáme s očakávaným vývojom na trhu práce. O vplyve jednotlivých vysvetľujúcich premenných na úspešnosť na trhu práce nám najviac prezradia hraničné vplyvy jednotlivých odhadnutých parametrov. Ďalej tiež uvedieme výsledky odhadu šancí pre Logit modely. Z hľadiska štatistickej významnosti bude dôležité sledovať p-hodnoty odhadov parametrov, hodnotu testu vierohodnostným pomerom každého modelu ako celku, McFadden-ovu štatistiku pseudo R^2 a mieru správnosti predikcie.

4.2 Výsledky odhadov parametrov

Výsledky odhadov parametrov metódou maximálnej vierohodnosti môžeme vidieť v tabuľke 4.1, ktorá zobrazuje aj vzťah medzi odhadmi parametrov Logit a Probit modelu. Ak prenásobíme odhady z Logit modelov $3^{1/2}/\pi$, dostaneme hodnoty porovnateľné s odhadmi z Probit modelov. P-hodnoty prislúchajúce k odhadnutým parametrom sú zachytené v tabuľke 4.2.

Tabuľka 4.1: Odhady parametrov a ich vzájomný vzťah v Logit a Probit modeli

Rok	2000			2006		
Model	Logit	$* 3^{1/2}/\pi$	Probit	Logit	$* 3^{1/2}/\pi$	Probit
Premenná	Koef.		Koef.	Koef.		Koef.
C	0,754898	0,416197	0,475474	0,461292	0,254324	0,343059
SEX	-0,032123	-0,017710	-0,021002	-0,360268	-0,198626	-0,207199
YOUNG	-0,488607	-0,269383	-0,273338	-0,300033	-0,165417	-0,160640
OLD	0,212048	0,116908	0,112281	0,070936	0,039109	0,015128
EDU	0,731476	0,403284	0,405325	1,085852	0,598662	0,570786
C_C	0,209779	0,115657	0,121393	0,398254	0,219569	0,214088
SINGLE	-0,737420	-0,406561	-0,420359	-0,544012	-0,299930	-0,294082
MARRIED	0,082321	0,045386	0,050185	0,230849	0,127274	0,135190
DIVORCED	-0,603581	-0,332772	-0,329892	-0,312032	-0,172032	-0,167707

Tabuľka 4.2: P-hodnoty odhadnutých parametrov v jednotlivých modeloch

Rok	2000		2006	
Model	Logit	Probit	Logit	Probit
Premenná	p-hodn.	p-hodn.	p-hodn.	p-hodn.
C	0,0000	0,0000	0,0000	0,0000
SEX	0,1565	0,1002	0,0000	0,0000
YOUNG	0,0000	0,0000	0,0000	0,0000
OLD	0,0000	0,0000	0,0527	0,4345
EDU	0,0000	0,0000	0,0000	0,0000
C_C	0,0000	0,0000	0,0000	0,0000
SINGLE	0,0000	0,0000	0,0000	0,0000
MARRIED	0,3534	0,3026	0,0087	0,0050
DIVORCED	0,0000	0,0000	0,0021	0,0025

Môžeme si všimnúť, že Logit aj Probit modely nám dávajú rovnaké výsledky smeru vplyvov pri štatisticky významných vysvetľujúcich premenných a to aj v oboch sledovaných rokoch, čo značí, že modelovaná situácia na trhu práce sa v priebehu šiestich rokov zásadne nezmenila.

V roku 2000 majú v Logit aj Probit modeli podľa očakávania negatívny vplyv na úspešnosť na trhu práce premenné SEX, YOUNG, SINGLE a DIVORCED. Ak respondenta predstavuje žena, mladý človek, slobodný alebo rozvedený, je nižšia pravdepodobnosť, že je zamestnaný. Predpovedaný pozitívny vplyv majú premenné EDU a MARRIED. Čím má človek vyšší stupeň vzdelania alebo je ženatý/vydatá, pravdepodobnosť, že je zamestnaný, je vyššia. Neočakávali sme pozitívny vplyv premenných OLD a C_C. Ak je človek starší alebo miesto jeho trvalého bydliska sa nachádza na vidieku, je vyššia pravdepodobnosť, že je zamestnaný.

Výsledky pozitívneho vplyvu premenných OLD, EDU, C_C, MARRIED a negatívneho vplyvu YOUNG, SINGLE a DIVORCED sa nám potvrdzujú aj z krížových tabuliek za rok 2000.³⁷ Smer vplyvu premennej SEX sa nedá jednoznačne charakterizovať.

Analogicky v roku 2006 vyšla negatívna závislosť úspešnosti na trhu práce od premenných SEX, YOUNG, SINGLE a DIVORCED. Ak respondent disponuje nasledovnými parametrami - žena, mladý človek, slobodný, rozvedený, pravdepodobnosť, že je zamestnaný, klesá. Na druhej

³⁷ Pozri prílohu 6.4.

strane, čím má vyšší stupeň vzdelania alebo je ženatý/vydatá, pravdepodobnosť je vyššia, čo svedčí o pozitívnej závislosti od premenných EDU a MARRIED. Podobne ako v roku 2000, ani v tomto období sme neočakávali pozitívny vplyv premenných OLD a C_C.

Výsledky pozitívneho vplyvu premenných EDU, C_C, MARRIED a negatívneho vplyvu YOUNG, SINGLE a DIVORCED sa nám potvrdzujú aj z krížových tabuliek za rok 2006.³⁸ Smer vplyvu premennej SEX a OLD sa nedá jednoznačne charakterizovať.

P-hodnota premennej SEX nadobúda v Logit modeli za rok 2000 hodnotu 0,1565 a v Probit modeli za ten istý rok 0,1002. V roku 2006 má v Logit modeli aj Probit modeli hodnotu menšiu ako 0,001. Veličina popisujúca rodovú príslušnosť sa v roku 2000 na hladine významnosti 5% ukázala štatisticky nevýznamná, čo môže byť spôsobené vyrovnanými hodnotami pozorovaní pre obe pohlavia, ako môžeme vidieť z krížovej tabuľky 6.5.

Výška p-hodnôt premennej MARRIED v Logit a Probit modeli za rok 2000 je rovná 0,3534 resp. 0,3026. V roku 2006 je však v oboch modeloch menšia ako 0,001, čo svedčí o jej štatistickej nevýznamnosti len pre rok 2000. Výsledky z krížovej tabuľky 6.11 však jasne hovoria, že skoro 87% ľudí v manželskom stave je zamestnaných.

Premenná OLD je štatisticky významná v roku 2000, lebo jej p-hodnota je na hladine významnosti 5% menšia ako 0,001 tak v Logit, ako aj v Probit modeli. V roku 2006 je však jej p-hodnota v Logit modeli na úrovni 0,0527 a v Probit modeli 0,4345. V tomto roku sa na hladine významnosti 5% v oboch modeloch javí ako štatisticky nevýznamná, čo môže byť spôsobené vyrovnanými hodnotami pozorovaní pre obe kategórie tejto premennej, ako môžeme vidieť z krížovej tabuľky 6.15.

Kvalitatívne znaky, predstavujúce pohlavie a manželský stav, nemali v roku 2000 vplyv na status respondenta na trhu práce. Rozlišovanie, či respondent patrí do vekovo staršej skupiny ľudí alebo do ďalších dvoch, nemalo význam pri skúmaní úspešnosti na trhu práce pre rok 2006.

Hraničné vplyvy jednotlivých odhadnutých parametrov modelov by mali mať najväčšiu výpovednú hodnotu vo vzťahu k úspešnosti na trhu práce. Ako však môžeme vidieť z tabuľky

³⁸ Pozri prílohu 6.4.

4.3, mnohé marginálne efekty vychádzajú menšie ako 0,1%. Takéto nízke hodnoty by nemalo význam interpretovať a nebudeme sa zaoberať ani štatisticky nevýznamnými premennými.

Podľa výsledkov z Logit modelu za rok 2000 môžeme vidieť, že mladý človek má o 0,22% vyššiu pravdepodobnosť, že je zamestnaný, ako ľudia z ostatných vekových skupín. Pozitívny hraničný vplyv je ale v rozpore s odhadnutým znamienkom parametra pri premennej YOUNG. Podobný výsledok nám ukazuje aj Probit model s tým rozdielom, že pravdepodobnosť je ešte nižšia (0,14%).

Za rok 2006 dostávame pre premennú YOUNG takisto rozporuplné výsledky v porovnaní so smerom vplyvu odhadnutého parametra. Hodnota pravdepodobnosti z Logit modelu je 1,14% a z Probit modelu 0,62%. V tomto roku má však hraničný vplyv premennej SEX zhodne s odhadnutým parametrom negatívny smer vplyvu. Teda žena má podľa výsledku z Logit modelu o 3,6% nižšiu pravdepodobnosť, že je zamestnaná, ako muž. V Probit modeli má táto pravdepodobnosť hodnotu -3,9%.

Hraničné vplyvy odhadov parametrov nám vo všeobecnosti nepriniesli očakávané informácie. Jedinou premennou, pre ktorú môžeme zmysluplne interpretovať marginálny efekt je SEX. Žena v porovnaní s mužom v roku 2006 má o 3,6%, resp. o 3,9% nižšiu pravdepodobnosť, že je úspešná na trhu práce, teda že má pracovné miesto.

Tabuľka 4.3: Hraničné vplyvy odhadnutých parametrov v jednotlivých modeloch

Rok	2000		2006	
Model	Logit	Probit	Logit	Probit
Premenná	hran.vplyv	hran.vplyv	hran.vplyv	hran.vplyv
SEX	-0,004377	-0,005066	-0,036291	-0,038688
YOUNG	0,002150	0,001385	0,011449	0,006215
OLD	0,000455	0,000155	0,000799	0,000094
EDU	0,000332	0,000063	0,000866	0,000054
C_C	0,000070	0,000008	0,000345	0,000011
SINGLE	-0,000051	-0,000003	-0,000187	-0,000003
MARRIED	-0,000004	0,000000	-0,000043	0,000000
DIVORCED	0,000003	0,000000	0,000013	0,000000

Na záver ešte uvádzame predpovedané hodnoty šancí z Logit modelu pre oba sledované roky. Pripomeňme, že šance sa vypočítajú ako pomer pravdepodobnosti výskytu želanej udalosti

v pomere s pravdepodobnosťou vyskytnutia sa neželanej udalosti. Takýmto spôsobom boli vypočítané šance pre všetky pozorovania a bol z nich spravený aritmetický priemer.

V roku 2000 vyšla hodnota šancí 7,429975, čo znamená, že šanca, že respondent je zamestnaný, je v priemere 7,5 ku 1. V roku 2006 sme vypočítali o niečo viac ako dvojnásobne vyššie šance. Hodnota 15,92213 vyjadruje šance, že v skoro 16-krát viac prípadoch bude účastník trhu práce zamestnaný.

4.3 Analýza priliehavosti modelov

Výsledky testovania mier priliehavosti pre Logit a Probit modely za roky 2000 a 2006 podrobne popisuje tabuľka 4.4.

Tabuľka 4.4: Miery priliehavosti modelov

Rok	2000		2006	
Model	Logit	Probit	Logit	Probit
% správnosti predikcie	0,820928	0,821079	0,873487	0,873375
<i>LR štatistika</i>	5088,535	5085,376	5548,394	5446,517
<i>P-hodnota LR</i>	0,0000	0,0000	0,0000	0,0000
<i>McFadden Pseudo R²</i>	0,089579	0,089523	0,132768	0,130330

Percentuálna predikcia statusu zamestnanosti ako vysvetľovanej premennej, vypočítaná na základe odhadov parametrov, nám ukazuje relatívne pozitívne výsledky. V roku 2000 tvoril pomer správne predpovedaných hodnôt pozorovaní nezávislej premennej v Logit aj Probit modeli niečo viac ako 82%. V roku 2006 bolo dokonca viac ako 87% hodnôt statusu respondenta na trhu práce určených správne, čo hovorí o pomerne dobrej predikčnej výpovednosti Logit a Probit modelov, vzhľadom na dané individuálne dáta. Podrobnejšie hodnoty sú zachytené v tabuľke 4.5.

Tabuľka 4.5: Predikcia statusu zamestnanosti v jednotlivých modeloch

Rok	2000			2006		
		Predikcia			Predikcia	
Hodnota	Dáta	Logit	Probit	Dáta	Logit	Probit
0	10920	631	572	7058	910	662
1	48771	59060	59119	46644	52792	53040
Spolu / %	59691	0,820928	0,821079	53702	0,873487	0,873375

Ďalším štatistickým ukazovateľom, vypovedajúcim o vhodnosti zvolených modelov pre dané dáta, je LR štatistika. Pomocou nej testujeme nulovú hypotézu, či sú koeficienty modelu, okrem konštanty, rovné nule. Ak má táto štatistika vysokú hodnotu, nulovú hypotézu zamietame.

Pre Logit a Probit model za rok 2000 má LR štatistika hodnotu 5088,535 resp. 5085,376. Nulovú hypotézu môžeme zamietnuť a tvrdíme, že nie všetky koeficienty sú nulové a modely sú preto štatisticky významné.

Pre sledovaný rok 2006 vyšla dokonca LR štatistika pre Logit a Probit model rovná hodnotám 5548,394, resp. 5446,517, čo je viac ako v roku 2000. Nulovú hypotézu, tvrdenie o nulovosti všetkých koeficientov a štatistickej nevýznamnosti modelov, môžeme preto opäť zamietnuť.

P-hodnota LR v Logit a Probit modeloch za sledované roky 2000 a 2006 vyšla menšia ako 0,0001. Pri nulovej hypotéze o koeficientoch modelu rovných nule má LR testovacia štatistika χ^2 rozdelenie so stupňami voľnosti rovnými počtu reštrikcií. V našom prípade s ôsmimi premennými je počet stupňov voľnosti rovný ôsmim. Kritická hodnota χ^2 rozdelenia s ôsmimi stupňami voľnosti na hladine významnosti 5% je 15,51, čo je viac ako p-hodnota vo všetkých skúmaných modeloch. Znamená to, že nie všetky koeficienty modelov sú nulové a samotné modely sú štatisticky významné.

Posledným štatistickým ukazovateľom vypovedajúcim o kvalite priliehavosti zvolených modelov k daným dátam je McFaddenova štatistika pseudo R^2 . Hodnota tejto štatistiky však vyšla za rok 2000 v oboch modeloch približne 0,09 a za rok 2006 len niečo vyše 0,13. Nízke hodnoty svedčia o tom, že nami definované regresné modely nemusia dostatočne dobre vystihovať dané individuálne dáta, a že skúmané vzťahy na trhu práce sú zrejme o niečo zložitejšie.

Záver

Modelovanie vplyvu kvalitatívnych znakov na úspešnosť na trhu práce pomocou metód maximálnej vierohodnosti sa ukázalo ako vcelku úspešné. Výsledky modelovania z Logit a Probit modelov za roky 2000 a 2006 boli veľmi podobné, čo potvrdzuje teoretický predpoklad podobnosti týchto dvoch modelov. V jednotlivých rokoch sa nám ukázal rovnaký smer vplyvu štatisticky významných vysvetľujúcich premenných na status respondenta na trhu práce. Situácia na trhu práce sa v priebehu šiestich rokov zásadne nezmenila.

Pre roky 2000 a 2006 môžeme spoločne povedať, že ak bol respondentom mladý človek, slobodný alebo rozvedený, bola nižšia pravdepodobnosť, že je zamestnaný. Ak sa miesto trvalého bydliska respondenta nachádzalo na vidieku alebo čím mal vyšší stupeň vzdelania, pravdepodobnosť, že je zamestnaný, bola vyššia. Tieto závislosti vyšli rovnaké pre oba sledované roky, preto je pravdepodobné, že situácia bude analogická aj pre obdobie medzi rokmi 2000 a 2006.

V roku 2000 sa navyše ukázalo, že ak človek patril do staršej vekovej kategórie (50 a viac rokov), bola nižšia pravdepodobnosť, že je zamestnaný. Pohlavie a skutočnosť či bol respondent ženatý/vydatá nemalo pre daný rok vplyv na status na trhu práce.

Ak v roku 2006 respondenta predstavovala žena, bola nižšia pravdepodobnosť, že je zamestnaná. Na druhej strane ak bol človek ženatý/vydatá, pravdepodobnosť, že je zamestnaný, bola vyššia. Či respondent patril do staršej vekovej kategórie (50 a viac rokov), nemalo pre daný rok vplyv na status na trhu práce.

Pre rok 2006 sa nám ďalej podarilo zistiť, že žena mala podľa výsledkov z Logit modelu o 3,6% nižšiu pravdepodobnosť, že je zamestnaná, ako muž. V Probit modeli mala táto negatívna pravdepodobnosť hodnotu 3,9%.

Z pohľadu celkovej zamestnanosti pre rok 2000 vyšla priemerná hodnota šancí, že respondent je zamestnaný, 7,5 ku 1. V roku 2006 sme vypočítali o niečo viac ako dvojnásobne vyššie šance. Skoro v 16-krát viac prípadoch bol účastník trhu práce zamestnaný ako nezamestnaný.

V roku 2000 tvoril pomer správne predpovedaných hodnôt pozorovaní nezávislej premennej v Logit aj Probit modeli niečo viac ako 82%. V roku 2006 bolo dokonca viac ako 87% hodnôt statusu respondenta na trhu práce určených správne, čo hovorí o pomerne dobrej predikčnej výpovednosti Logit a Probit modelov vzhľadom na dané individuálne dáta.

Zvolené Logit a Probit modely za roky 2000 a 2006 sa ukázali ako správne definované, lebo vyšli štatisticky významné. Výsledné hodnoty McFaddenovej štatistiky pseudo R^2 nám však naznačujú, že nami definované regresné modely nemusia dostatočne dobre vystihovať dané individuálne dáta. Skúmané vzťahy na trhu práce môžu byť o niečo zložitejšie, ako sme očakávali.

Tu zároveň vidíme priestor pre ďalšie možné pokračovanie práce. Modelovanie vplyvu kvalitatívnych znakov na úspešnosť na trhu práce by sa dalo rozšíriť o niektoré zaujímavé vysvetľujúce premenné. V prípade, ak by sa podarilo skúmať tranzíciu na trhu práce, by to mohli byť napr. premenné dĺžka nezamestnanosti a metóda hľadania práce.

5 Použitá literatúra a pramene

5.1 Použitá literatúra

1. Azudová, L.: *Postavenie domácností a vývoj sociálnej štruktúry v Slovenskej republike*, v Karol Morvay a kol.: Transformácia ekonomiky: Skúsenosti Slovenska, Ústav slovenskej a svetovej ekonomiky Slovenskej akadémie vied, Bratislava, 2005, ISBN 80-7144-143-0
2. Hatrák, M.: *Ekonometria*, Iura Edition, spol. s r.o., 211. publikácia edície Ekonómia, prvé vydanie, Bratislava, 2007, ISBN 978-80-8078-150-7
3. Kališová, L.: *Pomerové ukazovatele ako indikátory eventuálneho bankrotu podniku - ekonometrická analýza*, Bakalárska práca, Fakulta sociálnych vied, Inštitút ekonomických štúdií, Univerzita Karlova, Praha, 2004, s.3-6
4. Klimešová, A.: *Metódy klasifikácie pre posúdenie úveroschopnosti klienta*, Diplomová práca, Prírodovedecká fakulta, Masarykova Univerzita, Brno, 2007, s.10-25
5. Kotulič, R.: *Analýza spôsobov merania nezamestnanosti na Slovensku*, Zborník vedeckých prác katedry ekonómie a ekonomiky ANNO 2007, Prešovská Univerzita, Prešov, 2007, ISBN 978-80-8068-657-4
6. Kvetan, V., Palenik, V.: *Health and Morbidity in the Accession Countries*, Country Report - Slovak Republic, ENEPRI Research Report No. 30, Brussels, Belgium, 2006, ISBN 92-9074-648-0
7. Kvetan V., Páleník V., a spoluautori: *Analytický popis správania tranzitívnej ekonomiky (prípád Slovenskej republiky)*, v Karol Morvay a kol.: Transformácia ekonomiky: Skúsenosti Slovenska, Ústav slovenskej a svetovej ekonomiky Slovenskej akadémie vied, Bratislava, 2005, ISBN 80-7144-143-0
8. McFadden, D.L.: *Econometric Analysis of Qualitative Response Models*, v edícii Z. Griliches & M.D. Intriligator: Handbook of Econometrics, číslo 2, Elsevier, Amsterdam, North Holland, 1984, časť 6, kapitola 24, str. 1395-1457
9. Pindyck, R.S., Rubinfeld, D.L.: *Econometrics Models and Economic Forecast*, 4/e edícia, McGraw-Hill Higher Education, 2000, ISBN 0-07-118831-2, Part 2: Single-Equation Regression Models, Chapter 10: Nonlinear and maximum-likelihood estimation, Chapter 11: Models of qualitative choice

10. Senaj, M.: *Režimy výmenných kurzov v nových členských krajinách EÚ*, Diplomová práca, FMFI UK, Bratislava, 2005, s.22-23
11. Stankovičová, I.: *Logistická regresia a jej využitie v ekonomickej praxi*, Forum Statisticum Slovaca, SŠDS, Ročník III., číslo 1/2007, ISSN 1336-7420
12. Stankovičová, I. a Vojtková, M.: *Viacrozmerné štatistické metódy s aplikáciami*, Iura Edition, spol.s r.o., 231. publikácia edície Ekonómia, Bratislava, 2007, ISBN 978-80-8078-152-1
13. *Výberové zisťovanie pracovných síl (1994-2001)* – publikácia ŠÚ SR

5.2 Pramene

14. Croux, Ch.: *Binary Choice Models*, Lecture notes, Econometrics for business and industry 2007-2008, Faculty of Economics and Management, Leuven, Belgium
<http://www.econ.kuleuven.be/public/ndbae06/courses/busecon/logreg.pdf>
15. Fox, J.: *Logit and Probit Models*, Sociology 740, Lecture Notes 10, Department of Sociology, McMaster University, 2007
<http://socserv.mcmaster.ca/jfox/Courses/soc740/lecture-10-notes.pdf>
16. Leblang, D.: *Logit & Probit models*, Political Science 5108/7108, Advanced Data Analysis III, Department of Political Science, University of Colorado, 2004
http://polsci.colorado.edu/COURSES/PSCI_7108/lectures/Logit%20&%20Probit%20Models.pdf
17. Murphy, A.: *Regression with a Binary Dependent Variable - Logit & Probit Models*, Lecture notes, Econometrics, Hertford College, Oxford, 2007
<http://economics.hertford.ox.ac.uk/Anthony/Teaching/Econometrics/Probit%20and%20Logit.pdf>
18. Siebertová, Z.: *Modely s binárnou závislou premennou*, Ekonometria, Prednáška 11, FMFI UK, Bratislava, 2005
<http://www.defm.fmfi.uniba.sk/ludia/siebertova/ekonometria/lecture11.pdf>
19. Witkowski, V.: *Klasický lineárny model*, Ekonometria, Prednáška 3., FMFI UK, Bratislava, 2007
<http://aiolos.um.savba.sk/~viktor/Econ/Lecture3.pdf>

20. Witkovský, V.: *Maximum likelihood (ML)*, Ekonometria, Prednáška 5., FMFI UK, Bratislava, 2007
<http://aiolos.um.savba.sk/~viktor/Econ/Chapter5.pdf>
21. *Estimating Binary Models in EViews*, EViews 5 Help, Quantitative Micro Software, LLC, 2004
22. *Pseudo R-Squareds*, Introduction to SAS. UCLA: Academic Technology Services, Statistical Consulting Group
http://www.ats.ucla.edu/stat/mult_pkg/faq/general/Pseudo_RSquareds.htm (15.3.2008)

6 Prílohy

6.1 Odvodenie rovnice pre Logit model

$$\begin{aligned}\Lambda(z) &= \int_{-\infty}^z \lambda(x) dx = \int_{-\infty}^z \frac{e^{-x}}{(1 + e^{-x})^2} dx = \left[\frac{1}{1 + e^{-x}} \right]_{-\infty}^z = \frac{1}{1 + e^{-z}} - \frac{1}{1 + e^{\infty}} = \frac{1}{1 + e^{-z}} - 0 = \\ &= \frac{1}{1 + e^{-z}} = \frac{1}{1 + \frac{1}{e^z}} = \frac{1}{\frac{1 + e^z}{e^z}} = \frac{e^z}{1 + e^z}\end{aligned}$$

6.2 Odvodenie hraničného vplyvu pre Logit model

$$\begin{aligned}\frac{\partial P(y_i = 1)}{\partial X_i} &= \frac{\partial \Lambda(\hat{\beta} X_i)}{\partial X_i} = \left(\frac{1}{1 + e^{-(\hat{\beta} X_i)}} \right)' = \frac{0 - e^{-(\hat{\beta} X_i)}(-\hat{\beta})}{(1 + e^{-(\hat{\beta} X_i)})^2} = \frac{e^{-(\hat{\beta} X_i)}}{(1 + e^{-(\hat{\beta} X_i)})^2} \hat{\beta} = \\ &= \frac{1}{1 + e^{-(\hat{\beta} X_i)}} \left(\frac{e^{-(\hat{\beta} X_i)} + 1 - 1}{1 + e^{-(\hat{\beta} X_i)}} \right) \hat{\beta} = \frac{1}{1 + e^{-(\hat{\beta} X_i)}} \left(1 - \frac{1}{1 + e^{-(\hat{\beta} X_i)}} \right) \hat{\beta} = \Lambda(1 - \Lambda) \hat{\beta}\end{aligned}$$

6.3 Tabuľky odhadnutých modelov

Tabuľka 6.1: Logit model za rok 2000

Dependent Variable: ILO

Method: ML - Binary Logit (Quadratic hill climbing)

Date: 03/27/08 Time: 16:11

Sample: 1 59691

Included observations: 59691

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.754898	0.091707	8.231678	0.0000
SEX	-0.032123	0.022672	-1.416877	0.1565
YOUNG	-0.488607	0.029408	-16.61449	0.0000
OLD	0.212048	0.035427	5.985529	0.0000
EDU	0.731476	0.015341	47.68203	0.0000
C_C	0.209779	0.022721	9.232749	0.0000
SINGLE	-0.737420	0.092390	-7.981640	0.0000
MARRIED	0.082321	0.088708	0.927995	0.3534
DIVORCED	-0.603581	0.100284	-6.018694	0.0000
Mean dependent var	0.817058	S.D. dependent var		0.386622
S.E. of regression	0.369342	Akaike info criterion		0.866705
Sum squared resid	8141.429	Schwarz criterion		0.868061
Log likelihood	-25858.24	Hannan-Quinn criter.		0.867126
Restr. log likelihood	-28402.51	Avg. log likelihood		-0.433202
LR statistic (8 df)	5088.535	McFadden R-squared		0.089579
Probability(LR stat)	0.000000			
Obs with Dep=0	10920	Total obs		59691
Obs with Dep=1	48771			

Tabuľka 6.2: Probit model za rok 2000

Dependent Variable: ILO

Method: ML - Binary Probit (Quadratic hill climbing)

Date: 03/27/08 Time: 16:10

Sample: 1 59691

Included observations: 59691

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.475474	0.050627	9.391665	0.0000
SEX	-0.021002	0.012776	-1.643790	0.1002
YOUNG	-0.273338	0.016965	-16.11235	0.0000
OLD	0.112281	0.019123	5.871633	0.0000
EDU	0.405325	0.008449	47.97499	0.0000
C_C	0.121393	0.012895	9.413779	0.0000
SINGLE	-0.420359	0.051048	-8.234587	0.0000
MARRIED	0.050185	0.048681	1.030911	0.3026
DIVORCED	-0.329892	0.055717	-5.920844	0.0000
Mean dependent var	0.817058	S.D. dependent var	0.386622	
S.E. of regression	0.369342	Akaike info criterion	0.866758	
Sum squared resid	8141.448	Schwarz criterion	0.868114	
Log likelihood	-25859.82	Hannan-Quinn criter.	0.867179	
Restr. log likelihood	-28402.51	Avg. log likelihood	-0.433228	
LR statistic (8 df)	5085.376	McFadden R-squared	0.089523	
Probability(LR stat)	0.000000			
Obs with Dep=0	10920	Total obs	59691	
Obs with Dep=1	48771			

Tabuľka 6.3: Logit model za rok 2006

Dependent Variable: ILO

Method: ML - Binary Logit (Quadratic hill climbing)

Date: 03/27/08 Time: 16:14

Sample: 1 53702

Included observations: 53702

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.461292	0.092872	4.966973	0.0000
SEX	-0.360268	0.027949	-12.89030	0.0000
YOUNG	-0.300033	0.037725	-7.953107	0.0000
OLD	0.070936	0.036607	1.937777	0.0527
EDU	1.085852	0.018450	58.85309	0.0000
C_C	0.398254	0.027641	14.40821	0.0000
SINGLE	-0.544012	0.093442	-5.821936	0.0000
MARRIED	0.230849	0.087974	2.624068	0.0087
DIVORCED	-0.312032	0.101228	-3.082482	0.0021
Mean dependent var	0.868571	S.D. dependent var	0.337872	
S.E. of regression	0.315955	Akaike info criterion	0.675205	
Sum squared resid	5360.037	Schwarz criterion	0.676695	
Log likelihood	-18120.94	Hannan-Quinn criter.	0.675671	
Restr. log likelihood	-20895.14	Avg. log likelihood	-0.337435	
LR statistic (8 df)	5548.394	McFadden R-squared	0.132768	
Probability(LR stat)	0.000000			
Obs with Dep=0	7058	Total obs	53702	
Obs with Dep=1	46644			

Tabuľka 6.4: Probit model za rok 2006

Dependent Variable: ILO

Method: ML - Binary Probit (Quadratic hill climbing)

Date: 03/27/08 Time: 16:13

Sample: 1 53702

Included observations: 53702

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.343059	0.050926	6.736390	0.0000
SEX	-0.207199	0.015107	-13.71537	0.0000
YOUNG	-0.160640	0.020928	-7.675774	0.0000
OLD	0.015128	0.019356	0.781594	0.4345
EDU	0.570786	0.009734	58.64114	0.0000
C_C	0.214088	0.015075	14.20168	0.0000
SINGLE	-0.294082	0.051322	-5.730100	0.0000
MARRIED	0.135190	0.048182	2.805808	0.0050
DIVORCED	-0.167707	0.055453	-3.024305	0.0025
Mean dependent var	0.868571	S.D. dependent var	0.337872	
S.E. of regression	0.316524	Akaike info criterion	0.677102	
Sum squared resid	5379.381	Schwarz criterion	0.678592	
Log likelihood	-18171.88	Hannan-Quinn criter.	0.677568	
Restr. log likelihood	-20895.14	Avg. log likelihood	-0.338384	
LR statistic (8 df)	5446.517	McFadden R-squared	0.130330	
Probability(LR stat)	0.000000			
Obs with Dep=0	7058	Total obs	53702	
Obs with Dep=1	46644			

6.4 Krížové tabuľky

Tabuľka 6.5: Krížová tabuľka premenných ILO a SEX pre rok 2000

Tabulation of ILO and SEX

Date: 04/20/08 Time: 00:10

Sample: 1 59691

Included observations: 59691

Tabulation Summary

<u>Variable</u>	<u>Categories</u>		
ILO	2		
SEX	2		
Product of Categories	4		
Count			
% Row	SEX		
% Col	0	1	Total
0	5960	4960	10920
	54.58	45.42	100.00
	18.87	17.64	18.29
ILO	25621	23150	48771
	52.53	47.47	100.00
	81.13	82.36	81.71
Total	31581	28110	59691
	52.91	47.09	100.00
	100.00	100.00	100.00

Tabuľka 6.6: Krížová tabuľka premenných ILO a YOUNG pre rok 2000

Tabulation of ILO and YOUNG

Date: 04/20/08 Time: 00:10

Sample: 1 59691

Included observations: 59691

Tabulation Summary

<u>Variable</u>		<u>Categories</u>		
ILO		2		
YOUNG		2		
Product of Categories		4		

Count				
% Row		YOUNG		
% Col		0	1	Total
0		6075	4845	10920
		55.63	44.37	100.00
		14.36	27.88	18.29
ILO 1		36237	12534	48771
		74.30	25.70	100.00
		85.64	72.12	81.71
Total		42312	17379	59691
		70.89	29.11	100.00
		100.00	100.00	100.00

Tabuľka 6.7: Krížová tabuľka premenných ILO a OLD pre rok 2000

Tabulation of ILO and OLD

Date: 04/20/08 Time: 00:11

Sample: 1 59691

Included observations: 59691

Tabulation Summary

<u>Variable</u>	<u>Categories</u>		
ILO	2		
OLD	2		
Product of Categories	4		
Count			
% Row	OLD		
% Col	0	1	Total
0	9663	1257	10920
	88.49	11.51	100.00
	19.55	12.26	18.29
ILO	39776	8995	48771
	81.56	18.44	100.00
	80.45	87.74	81.71
Total	49439	10252	59691
	82.82	17.18	100.00
	100.00	100.00	100.00

Tabuľka 6.8: Krížová tabuľka premenných ILO a EDU pre rok 2000

Tabulation of ILO and EDU

Date: 04/20/08 Time: 00:12

Sample: 1 59691

Included observations: 59691

Tabulation Summary

<u>Variable</u>		<u>Categories</u>				
ILO		2				
EDU		4				
Product of Categories		8				
Count						
% Row		EDU				
% Col		0	1	2	3	Total
0		2252	4833	3527	308	10920
		20.62	44.26	32.30	2.82	100.00
		38.12	20.09	14.70	5.37	18.29
ILO 1		3656	19227	20461	5427	48771
		7.50	39.42	41.95	11.13	100.00
		61.88	79.91	85.30	94.63	81.71
Total		5908	24060	23988	5735	59691
		9.90	40.31	40.19	9.61	100.00
		100.00	100.00	100.00	100.00	100.00

Tabuľka 6.9: Krížová tabuľka premenných ILO a C_C pre rok 2000

Tabulation of ILO and C_C

Date: 04/20/08 Time: 00:12

Sample: 1 59691

Included observations: 59691

Tabulation Summary

<u>Variable</u>	<u>Categories</u>		
ILO	2		
C_C	2		
Product of Categories	4		
Count			
% Row		C_C	
% Col	0	1	Total
0	4987	5933	10920
	45.67	54.33	100.00
	22.38	15.86	18.29
ILO	17295	31476	48771
	35.46	64.54	100.00
	77.62	84.14	81.71
Total	22282	37409	59691
	37.33	62.67	100.00
	100.00	100.00	100.00

Tabuľka 6.10: Krížová tabuľka premenných ILO a SINGLE pre rok 2000

Tabulation of ILO and SINGLE

Date: 04/20/08 Time: 00:12

Sample: 1 59691

Included observations: 59691

Tabulation Summary

<u>Variable</u>		<u>Categories</u>		
ILO		2		
SINGLE		2		
Product of Categories		4		

Count				
% Row		SINGLE		
% Col		0	1	Total
0		6057	4863	10920
		55.47	44.53	100.00
		13.89	30.25	18.29
ILO 1		37558	11213	48771
		77.01	22.99	100.00
		86.11	69.75	81.71
Total		43615	16076	59691
		73.07	26.93	100.00
		100.00	100.00	100.00

Tabuľka 6.11: Krížová tabuľka premenných ILO a MARRIED pre rok 2000

Tabulation of ILO and MARRIED

Date: 04/20/08 Time: 00:13

Sample: 1 59691

Included observations: 59691

Tabulation Summary

<u>Variable</u>		<u>Categories</u>		
ILO		2		
MARRIED		2		
Product of Categories		4		

Count				
% Row		MARRIED		
% Col		0	1	Total
0		5586	5334	10920
		51.15	48.85	100.00
		28.27	13.36	18.29
ILO 1		14176	34595	48771
		29.07	70.93	100.00
		71.73	86.64	81.71
Total		19762	39929	59691
		33.11	66.89	100.00
		100.00	100.00	100.00

Tabuľka 6.12: Krížová tabuľka premenných ILO a DIVORCED pre rok 2000

Tabulation of ILO and DIVORCED

Date: 04/20/08 Time: 00:13

Sample: 1 59691

Included observations: 59691

Tabulation Summary

<u>Variable</u>		<u>Categories</u>		
ILO		2		
DIVORCED		2		
Product of Categories		4		

Count				
% Row		DIVORCED		
% Col		0	1	Total
0		10365	555	10920
		94.92	5.08	100.00
		18.15	21.61	18.29
ILO 1		46758	2013	48771
		95.87	4.13	100.00
		81.85	78.39	81.71
Total		57123	2568	59691
		95.70	4.30	100.00
		100.00	100.00	100.00

Tabuľka 6.13: Krížová tabuľka premenných ILO a SEX pre rok 2006

Tabulation of ILO and SEX

Date: 04/20/08 Time: 16:23

Sample: 1 53702

Included observations: 53702

Tabulation Summary

<u>Variable</u>		<u>Categories</u>		
ILO		2		
SEX		2		
Product of Categories		4		

Count				
% Row		SEX		
% Col		0	1	Total
0		3556	3502	7058
		50.38	49.62	100.00
		12.29	14.14	13.14
ILO 1		25374	21270	46644
		54.40	45.60	100.00
		87.71	85.86	86.86
Total		28930	24772	53702
		53.87	46.13	100.00
		100.00	100.00	100.00

Tabuľka 6.14: Krížová tabuľka premenných ILO a YOUNG pre rok 2006

Tabulation of ILO and YOUNG

Date: 04/20/08 Time: 16:23

Sample: 1 53702

Included observations: 53702

Tabulation Summary

<u>Variable</u>	<u>Categories</u>
ILO	2
YOUNG	2
Product of Categories	4

Count			
% Row	YOUNG		
% Col	0	1	Total
0	4667	2391	7058
	66.12	33.88	100.00
	11.47	18.38	13.14
ILO	36025	10619	46644
	77.23	22.77	100.00
	88.53	81.62	86.86
Total	40692	13010	53702
	75.77	24.23	100.00
	100.00	100.00	100.00

Tabuľka 6.15: Krížová tabuľka premenných ILO a OLD pre rok 2006

Tabulation of ILO and OLD

Date: 04/20/08 Time: 16:24

Sample: 1 53702

Included observations: 53702

Tabulation Summary

<u>Variable</u>	<u>Categories</u>		
ILO	2		
OLD	2		
Product of Categories	4		
Count			
% Row	OLD		
% Col	0	1	Total
0	5638	1420	7058
	79.88	20.12	100.00
	13.75	11.18	13.14
ILO	35365	11279	46644
	75.82	24.18	100.00
	86.25	88.82	86.86
Total	41003	12699	53702
	76.35	23.65	100.00
	100.00	100.00	100.00

Tabuľka 6.16: Krížová tabuľka premenných ILO a EDU pre rok 2006

Tabulation of ILO and EDU

Date: 04/20/08 Time: 16:24

Sample: 1 53702

Included observations: 53702

Tabulation Summary

<u>Variable</u>	<u>Categories</u>				
ILO	2				
EDU	4				
Product of Categories	8				
Count					
% Row	EDU				
% Col	0	1	2	3	Total
0	2067	2804	1945	242	7058
	29.29	39.73	27.56	3.43	100.00
	47.79	14.82	8.52	3.18	13.14
ILO	2258	16122	20894	7370	46644
	4.84	34.56	44.79	15.80	100.00
	52.21	85.18	91.48	96.82	86.86
Total	4325	18926	22839	7612	53702
	8.05	35.24	42.53	14.17	100.00
	100.00	100.00	100.00	100.00	100.00

Tabuľka 6.17: Krížová tabuľka premenných ILO a C_C pre rok 2006

Tabulation of ILO and C_C

Date: 04/20/08 Time: 16:24

Sample: 1 53702

Included observations: 53702

Tabulation Summary

<u>Variable</u>		<u>Categories</u>		
ILO		2		
C_C		2		
Product of Categories		4		
Count				
% Row		C_C		
% Col		0	1	Total
0		3494	3564	7058
		49.50	50.50	100.00
		18.35	10.28	13.14
ILO 1		15544	31100	46644
		33.32	66.68	100.00
		81.65	89.72	86.86
Total		19038	34664	53702
		35.45	64.55	100.00
		100.00	100.00	100.00

Tabuľka 6.18: Krížová tabuľka premenných ILO a SINGLE pre rok 2006

Tabulation of ILO and SINGLE

Date: 04/20/08 Time: 16:25

Sample: 1 53702

Included observations: 53702

Tabulation Summary

<u>Variable</u>		<u>Categories</u>		
ILO		2		
SINGLE		2		
Product of Categories		4		

Count				
% Row		SINGLE		
% Col		0	1	Total
0		4178	2880	7058
		59.20	40.80	100.00
		10.74	19.48	13.14
ILO 1		34740	11904	46644
		74.48	25.52	100.00
		89.26	80.52	86.86
Total		38918	14784	53702
		72.47	27.53	100.00
		100.00	100.00	100.00

Tabuľka 6.19: Krížová tabuľka premenných ILO a MARRIED pre rok 2006

Tabulation of ILO and MARRIED

Date: 04/20/08 Time: 16:25

Sample: 1 53702

Included observations: 53702

Tabulation Summary

<u>Variable</u>	<u>Categories</u>		
ILO	2		
MARRIED	2		
Product of Categories	4		

Count			
% Row	MARRIED		
% Col	0	1	Total
0	3535	3523	7058
	50.09	49.91	100.00
	18.80	10.10	13.14
ILO	15269	31375	46644
	32.74	67.26	100.00
	81.20	89.90	86.86
Total	18804	34898	53702
	35.02	64.98	100.00
	100.00	100.00	100.00

Tabuľka 6.20: Krížová tabuľka premenných ILO a DIVORCED pre rok 2006

Tabulation of ILO and DIVORCED

Date: 04/20/08 Time: 16:26

Sample: 1 53702

Included observations: 53702

Tabulation Summary

<u>Variable</u>	<u>Categories</u>
ILO	2
DIVORCED	2
Product of Categories	4

Count	DIVORCED		
% Row	0	1	Total
% Col	0	1	Total
0	6595	463	7058
	93.44	6.56	100.00
	12.97	16.17	13.14
ILO	44243	2401	46644
	94.85	5.15	100.00
	87.03	83.83	86.86
Total	50838	2864	53702
	94.67	5.33	100.00
	100.00	100.00	100.00