Warsaw School of Economics
Institute of Econometrics
Department of Applied Econometrics

Department of Applied Econometrics Working Papers
Warsaw School of Economics
Al. Niepodleglosci 164
02-554 Warszawa, Poland

Working Paper No. 2-11

Determinants of involuntary job termination in the Polish labor market

Krzysztof Pytka
Warsaw School of Economics (SGH), Collegium Invisibile

This paper is available at the Warsaw School of Economics
Department of Applied Econometrics website at: http://www.sgh.waw.pl/instytuty/zes/wp/

Determinants of Involuntary Job Termination in the Polish Labor Market

Krzysztof Pytka*

April 11, 2011

Abstract

The intention of this article is to evaluate the exogenous dismissal probability for a certain worker depending on her characteristics for the Polish labor market. To model this phenomenon I considered a range of count data models. In the analysis the data from the Polish General Social Survey of 2008 was used. Covariates explaining a number of unemployment spells were selected in the spirit of the human-capital theory. In the course of the study existence of intransferable firm-specific human capital across employers and depreciation of the human capital acquired through learning by doing have been empirically confirmed. The conducted analysis may be considered the first step in the calibration of a job-search model with heterogeneous agents.

Keywords: Involuntary Unemployment, Count Data Models, Job Search

JEL Code: J63, J64

*Warsaw School of Economics (SGH), Collegium Invisibile; e-mail:K.Pytka@ci.edu.pl

1 Motivation

Everyone is one-of-a-kind. It may sound like a very bad cliché in the real world but in the contemporary economic theory the problem of heterogeneity amongst agents interests an increasing number of economists. Theoretic models augmented with heterogeneous agents may offer new qualitative conclusions (e.g. Burdett and Mortensen, 1998) and may describe real economies more precisely. This remark concerns job-search models as well. Job-search models are an important part of state-of-the-art labor economics and are commonly used nowadays.

The intention of this article is to furnish with the exogenous dismissal probability for a certain worker depending on her characteristics for the Polish labor market. The job-separation probability is one of the key elements in the fundamental reservation wage equation (Burdett and Mortensen, 1980). The estimation of this parameter would pose a contribution into calibration of the job-search model with heterogeneous workers for the Polish economy.

In the course of the analysis I decided to model the dismissal probability in the spirit of the human capital theory (Mincer, 1962; Becker, 1962). Initially, this theory was developed to explain salary dispersion between workers. Nonetheless, the probability of being laid off can be perceived as a sort of nonpecuniary fringe benefits, which (like wages) depend on the accumulated human capital of a certain employer. Therefore, in the study such variables as overall experience, tenure at the current/last workplace, or years of education were selected to be explanatory.

To model the examined phenomenon I decided to use count-data models. This is motivated by the available data, which is from Polish General Social Survey of 2008 (Cichomski, Jerzy, and Zielinski, 2009). Admittedly, duration models are more common for this purpose. However, both classes of models are equivalent to each other (Winkelmann and Boes, 2006).

The plan of the paper is as follows. In section 2 there is a description of the used data set. In section 3 an estimation strategy is laid out briefly. Section 4 is dedicated to presentation of the empirical results. In the last section I outlined the conclusions of the conducted analysis.

2 Data Description

In the course of the analysis I used data from the Polish General Social Survey of 2008 (Cichomski, Jerzy, and Zielinski, 2009). The survey is the statutory research of the Institute for Social Studies, Warsaw University (ISS UW). Its structure is integrated with International Social Survey Programme and partly consistent with similar surveys conducted in other countries, such as General Social Survey (GSS, National Opinion Research Center, University of Chicago) and a German program of national social surveys (ALLBUS, Zentrum fur Umfragen, Methoden und Analysen, Mannheim).

I included information on all respondents who were employed at least for one month in the last 10 years and belonged to the total labor force while the survey was made. The first restriction is motivated by the obvious fact that a person unemployed through the whole period of the observation cannot be laid off. The latter is imposed by a character of the collected data. The survey provides information only on (total) unemployment duration and dropping all observations outside the work force while the survey I reduced a probability of including units who are not actively participating in the labor market. A time scope of the sample is limited to 1997–2008. For this period all the data used was available. Besides, a deeply transient character of the Polish economy in the first half of the 1990's seems to constitute an ample premise to exclude this period from the study as well. In the sample there were eventually 3,156 observations.

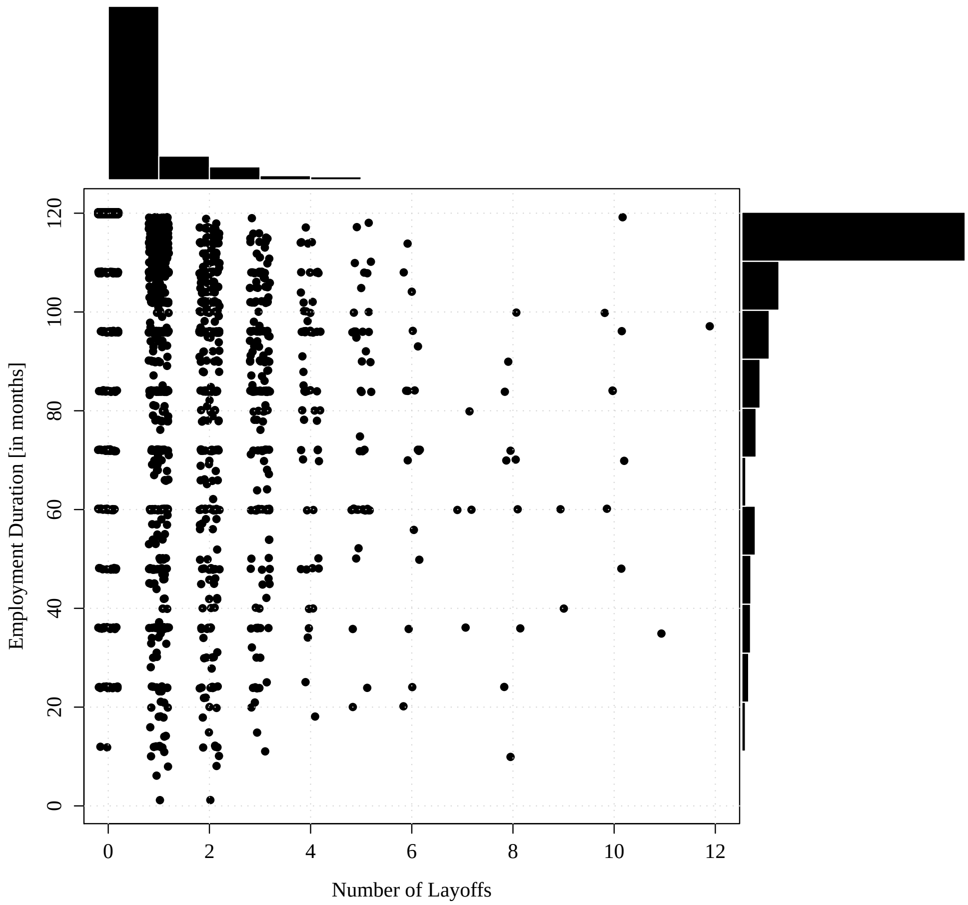
A number of unemployment periods is the dependent variable, which embodies involuntary job separations. This is quite a standard manner for measurement of this phenomenon (e.g. Börsch-Supan, 1990; Winkelmann and Zimmermann, 1998; Topel and Ward, 1992). Fig. 1 presents histograms of the examined variable, employment duration and a relation between them. The importance of employment duration for the analysis was explained in section 3.

Table 2 presents a set of covariates used in the study. Years of education in the model aim at seizing a level of human capital accumulated through Mincerian investments. Experience at the current workplace for the employed and at the last workplace for the unemployed is to control human capital acquired through *learning-*

Table 1: Unemployment spells amongst pollees from 1997 through 2008

Unemployment spells	1997	1999	2002	2005	2008
0	506	410	445	213	221
1	160	145	216	87	80
2	67	78	99	60	36
3	30	36	59	30	31
4	9	10	25	10	6
5	3	8	12	8	13
6	2	4	4	3	4
7	0	1	1	1	1
8	0	2	4	1	3
9	0	0	1	1	0
10	0	1	5	1	1
11	0	0	0	1	0
12	0	0	1	0	0

Figure 1: Unemployment spells number histogram vs. employment duration histogram



Source: Own computation

by-doing. I decided to distinguish last job experience between the employed and unemployed so as to measure *firm-specific* human capital, which is not transferable across employers. Apart from the mentioned covariates dummy variables identifying gender and voivodeships were included in the study.

Table 2: Variable descriptions

Variable	Description
exp	Overall work experience since 14 [in years].
$tenure$	Experience at the current workplace [in years, employed persons].
$tenure\_un$	Experience at the last workplace [in years, unemployed persons].
yrs_edu	Education [in years].
$female$	Dummy variable equal to 1 for women.
t	Employment duration in the last 10 years [in months].
$voiev16$	Set of dummies identifying one of 16 voivodeships. Reference category: the country mean in 1997 (before 1999 Administration Reform).

3 Estimation Strategy

In the literature of empirical labor economics duration models are used traditionally for calibrating job-search models (French and Tabe, 2010; Eckstein and van den Berg, 2007; Flinn, 2002). Nonetheless, because the data did not enable to conduct such an analysis I decided to estimate a count-data model instead. As Winkelmann and Boes (2006, p. 284) point out, in fact, both classes of models are actually equivalent to each other. They present the equation for the identity relationship between count-data and duration models. Both approaches are therefore equally correct to estimate individual job destruction probability.

Amongst count-data models the Poisson regression model is a benchmark for further improvements. To be eligible for using this model the following assumptions on the data generating process of the unemployment spells have to be made:

1. the proportionality of the likelihood for at least one occurrence of the event to the length of the time interval;

2. the probability of two or more occurrences can be neglected;

3. the numbers of occurrences in disjoint time intervals are independent.

For these conditions a number of occurrences in a fixed time interval is Poisson and its mean value y conditioned by covariates $\mathbf{x}$ can be conveyed by:

$$\mathbb{E}(y|\mathbf{x}) = \exp(\beta'\mathbf{x}). \tag{1}$$

The first complication in the study was that respondents varied in employment duration (t). The standard Poisson models assume that the risk period is the same for all observations. It is quite obvious that an individual who was relatively longer employed had more ‘opportunities’ to be laid off than, for instance, an unemloyed one, whose dismissal probability amounts to simply zero. The problem of different exposure durations can be tackled by modelling a mean value per time interval, *viz.*

$$\frac{\mathbb{E}(y|\mathbf{x})}{t} = \exp(\beta'\mathbf{x}). \tag{2}$$

Following (Winkelmann, 2003) I was allowed to make this transformation on the basis of the assumption (1). At the stage of estimation I rearranged the formula to the following one:

$$\mathbb{E}(y|\mathbf{x}) = (\beta'\mathbf{x} + \gamma \ln t). \tag{3}$$

Both equations are equivalent to each other if a restriction $\gamma = 1$ is made. However, as it is seen in fig. 1 a relationship between a number of unemployment spells and employment duration is ambiguous. As mentioned before, a longer employment duration means more ‘opportunities’ to be laid off. On the other hand, an individual who is longer employed posses relatively more information on the labor market and she is better accommodated to any changes at the workplace. Therefore, it seems prudent not to make a restriction on γ to verify whether there is an additional impact of t other than a logarithmic exposure offset.

One of features of the Poisson distribution is equidispersion, which concerns in the equality of both the mean value and the variance, i.e. $\mathbb{E}(y|\mathbf{x}) = Var(y|\mathbf{x})$. In the literature (Greene, 2008) it is stated that there may be two main sources of failure to meet this assumption: unobserved heterogeneity or an excess number of zeros.

If the first source of overdispersion occurs it can be solved by using NegBin models. It concerns in estimating one additional parameter accounting for the difference between the conditional expected value $\mathbb{E}(y|\mathbf{x})$ and the conditional variance

$Var(y|\mathbf{x})$. There are two variants of the model - NegBin I and NegBin II. In the course of the analysis I considered only NegBin II for its greater popularity (Greene, 2005). In this model the additional parameter θ is estimated for the following equation $Var(y|\mathbf{x}) = \exp(\beta'\mathbf{x}) + \theta^2[\exp(\beta'\mathbf{x})]^2$.

An excess number of zeros may be another source of the overdispersion. If it occurs there is premise to have an inkling that the dependent variable is generated by a mixture of distributions. Lambert (1992) proposed a solution to this problem. In her zero-inflated Poisson model (ZIP) there are two data generating processes: one accounting for zeros values and a second accounting for a number of positive occurrences. In other words, she states that the variable is censored, i.e.:

$$y = \begin{cases} 0 & \text{if } c = 1 \\ y^* & \text{if } c = 0 \end{cases} \tag{4}$$

Let ω be the probability of $c = 1$, which is described by a binomial choice model. Then the probability function $f(\cdot)$ of y can be presented as follows:

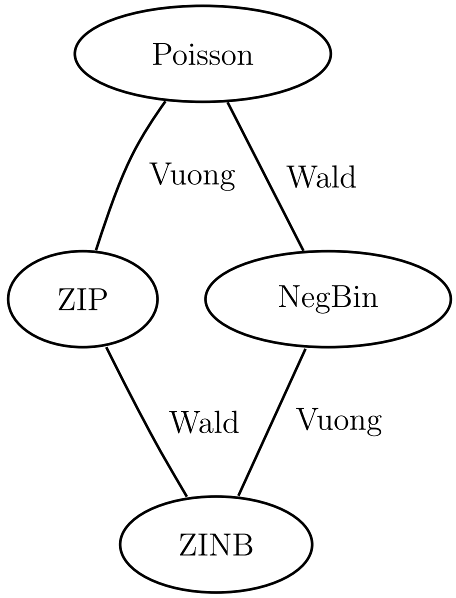
$$f(y) = \omega d + (1 - \omega)g(y), \tag{5}$$

where $d = 1 - \min\{1, y\}$ and $g(\cdot)$ is the probability function estimated from the Poisson regression model for individuals with a non-zero number of unemployment spells.

It is worth noting that one source of overdispersion does not exclude the other and both may coincide. In such a case Greene (1994) proposed to use the zero-inflated NegBin model (ZINB), which differs from the ZIP model with modelling the function $g(\cdot)$ by the use of NegBin regression.

The Poisson regression and the ZIP model are nested against the NegBin model and the ZINB model, respectively. Therefore to infer on unobserved heterogeneity causing overdispersion I am allowed to use the Wald test. Nonetheless, Greene (1994, p. 15-16) stresses out that the Poisson and the ZIP model are not nested against the NegBin. To choose between non-nested models I used Vuong test (Vuong, 1989). The statistical inference procedure is depicted in fig. 2.

Figure 2: Statistical inference procedure for count-data models



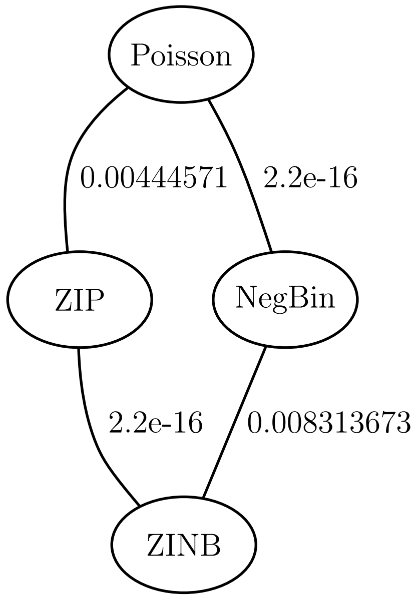
4 Empirical Results

In the course of the analysis I considered the following models: the Poisson model, the NegBin II model, the ZIP model, and the ZINB model. I used statistical tests mentioned before to infer which of them describes the examined phenomenon the best. As fig. 3 shows, on the basis of the conducted tests I ascertained that the ZINB is the best to explain the number of unemployment spells.

Tables 3 and 4 show the results of the estimation. Apart from one regional dummy and a covariate denoting tenure of an unemployed individual at the last workplace all variables are statistically significant at $\alpha = 0.05$. Moreover, an additional Wald test with a linear restriction $\beta_{tenure} = \beta_{tenure\_un}$ proved that both parameters are statistically different. This observation confirms the theory that a part of the human capital acquired through learning-by-doing cannot be transferred across employers and is only useful for the particular firm.

Signs of estimates for overall experience (linear and squared variables) in the count equation (table 4) may be surprising at first since they are at odds with - to some extent - similar studies on labor mobility (Winkelmann and Zimmermann,

Figure 3: Statistical inference results (p-value on arches).



1998; Winkelmann, 2003). In the cited works the linear term estimate exhibits a negative sign and the squared part is positive while in my results signs are opposite. Such a state of affairs can be explained by the set of the remaining covariates. In my analysis there are covariates of tenure (for the employed and the unemployed) and employment duration in the last 10 years. In fact, a whole impact of overall experience associated with human capital accumulation was seized by these 3 other covariates. What remains in the influence of the variable was linked only with age. This remark shows a phenomenon of depreciation of the human capital (in our analysis with a decreasing pace) and the predominant role of the 'up-to-date' experience.

Logarithm of employment duration in the last ten years was introduced into the count equation of the model so that all observations could be comparable. If the exposure effect existed only, the estimate of γ should be equal to 1. In my results $\hat{\gamma} = -0.4101964$ and differs from unit considerably. In fact we have two counteracting effects of employment duration:

$$\hat{\gamma} = \hat{\eta} + 1 = -0.4101964. \tag{6}$$

In the above equation we have decomposition of both effects. Unit stands for the exposure effect and plays a role of a logarithmic offset. Effect $\hat{\eta} = -1.41$ is an additional effect of employment duration which is connected with the mentioned 'up-to-date' experience. Its level surpasses the exposure effect and eventually the resultant impact of the both is negative.

Concerning the selection equation (table 3) I was made to introduce the logarithm offset for the unemployment period with a restriction. This was necessary for achieving the MLE convergence. The goal of the selection equation is to distinguish zero and positive-count observations. The overwhelming majority (91%) of individuals of the count of zero exhibit employment duration of 120 months and, obviously, there are no individuals with employment duration of 120 months with a non-zero number of the unemployment spells (see fig. 1). This implied that including the exposure effect without any restrictions would result in divergence of the algorithm. Unfortunately, imposing the restriction $\alpha_t = 1$ in the selection equation may have led to seizing the influence of the employment duration by the covariates accounting for overall experience. As a result, there is a discrepancy in signs of estimates for this variable between both equations.

In the model I included dummies denoting voivodeships comparing the mean 1997 country level. Fig. 4 presents a multiplying influence of the occupation place on the dismissal probability *ceteris paribus*. This multiplier is quite scattered. The probability of dismissal is nearly twice as big in the Podkarpackie region as in Małopolskie *ceteris paribus*.

The comparison of the goodness of fit is presented in fig. 5 in the manner proposed by (Long, 1997). It illustrates differences in fractions of predicted counts between the model and the observed values. We can see that there is small improvement in prediction performance of the Poisson (red line) and the ZIP model (green line) comparing the Poisson model with constant only (yellow line). Admittedly, in spite of the result of the Vuong test the selection equation in the ZIP model does not better a predictive quality of the model. However, this equation brought amelioration in the ZINB model comparing the NegBin II model. The binomial equation slightly underestimates the zero fraction. Eventually, in the ZINB model the maximal difference between the theoretical fraction and the empirical fraction is equal to 0.039 for the count of 1.

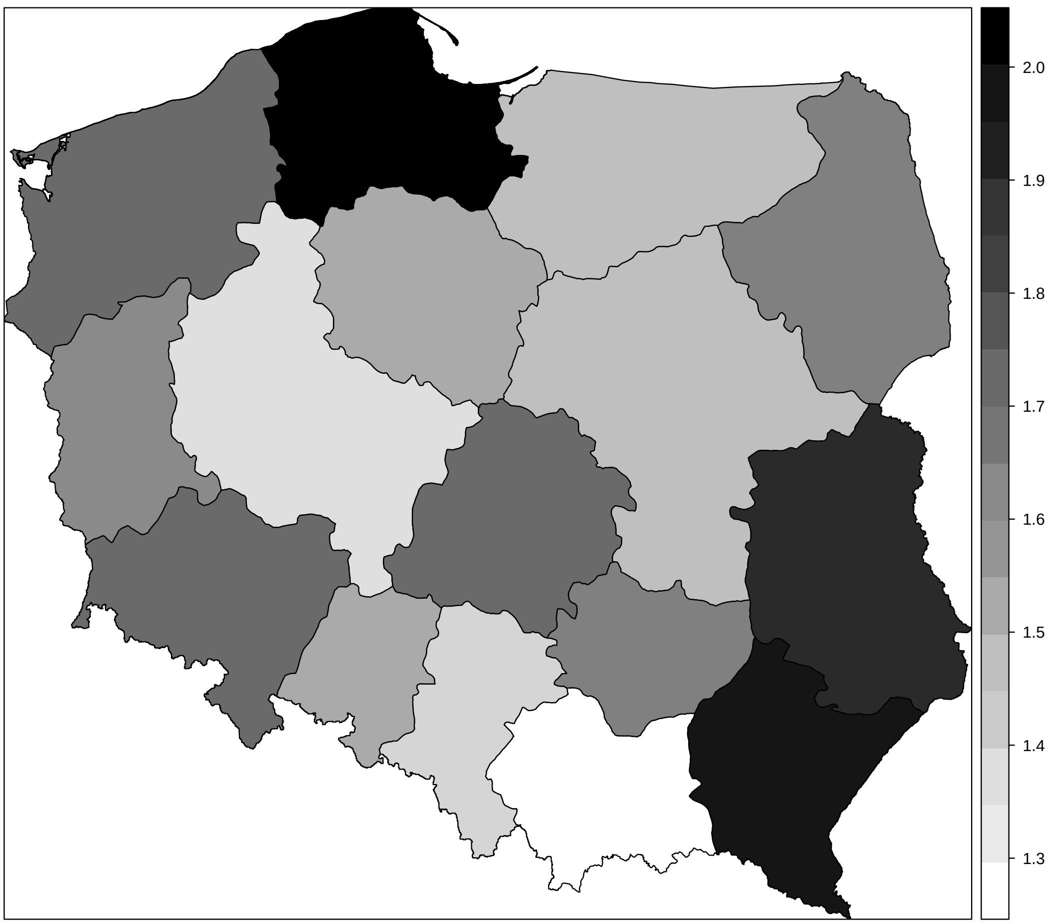
Table 3: Zero-inflated negative-binomial regression results for a number of layoffs. Selection equation (logit link) with the logarithmic offset of t.

Variable	Estimate	Std. Error	p-value
(Intercept)	-9.551547	1.180332	5.86E-016
exp	0.484058	0.098179	8.21E-007
exp^2	-0.010223	0.002212	3.82E-006
$tenure$	0.221256	0.01963	$<$ 2e-16
age	-0.110921	0.035598	0.00183
yrs_edu	0.169652	0.053633	0.00156

Table 4: Zero-inflated negative-binomial regression results for the number of layoffs. Count equation.

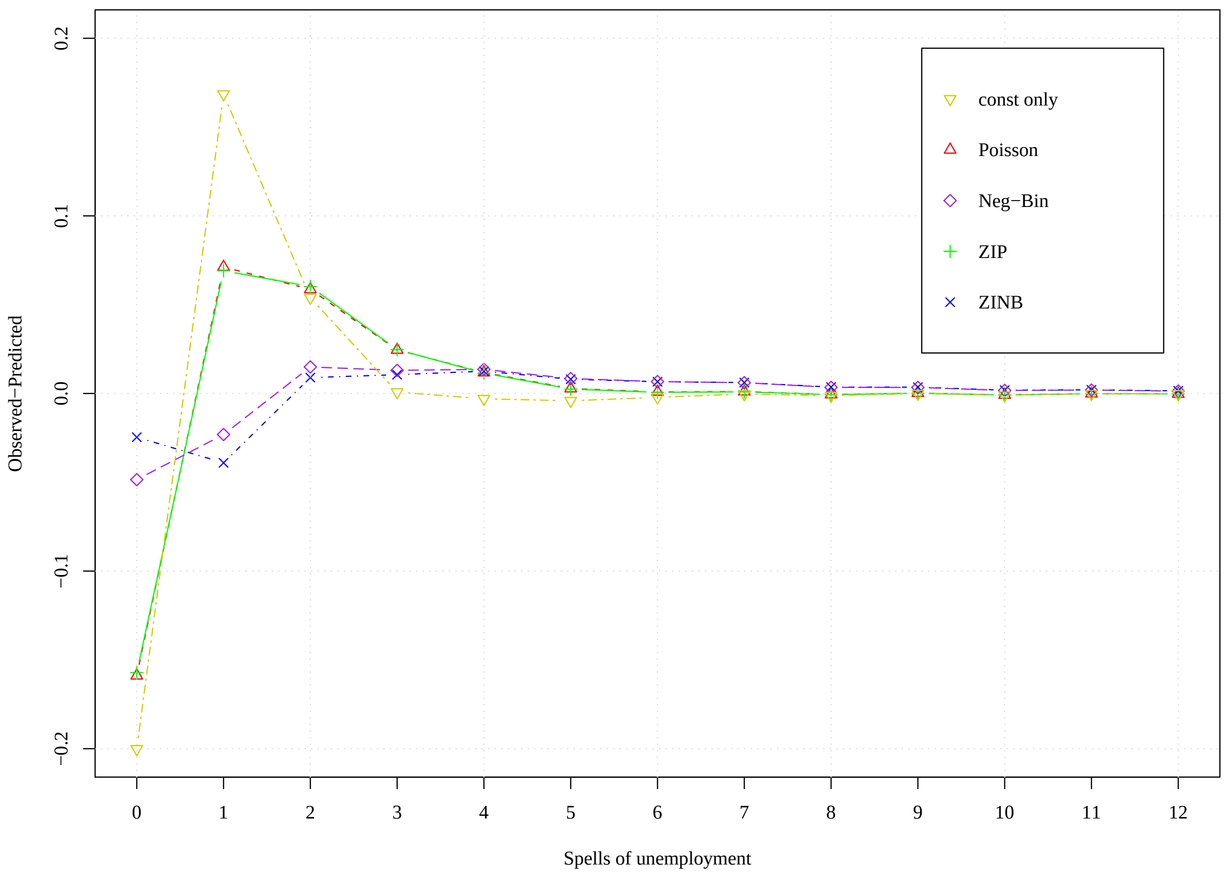
Variable	Estimate	Std. Error	p-value
(Intercept)	2.4968105	0.2646474	$<$ 2e-16
Dolnośląskie	0.5252591	0.1007174	1.84E-007
Kujawsko-Pomorskie	0.4352118	0.1198918	0.000283
Lubelskie	0.6568332	0.1391172	2.34E-006
Lubuskie	0.4636921	0.1371465	0.000722
Łódzkie	0.5264386	0.1064458	7.59E-007
Małopolskie	0.2613364	0.1390907	0.060259
Mazowieckie	0.3715494	0.0897977	3.51E-005
Opolskie	0.43533	0.172861	0.01179
Podkarpackie	0.6567169	0.1219539	7.25E-008
Podlaskie	0.5087641	0.1723553	0.003159
Pomorskie	0.6846627	0.115553	3.12E-009
Śląskie	0.3270036	0.092191	0.00039
Świętokrzyskie	0.5344152	0.1686163	0.001527
Warmińsko-Mazurskie	0.3933553	0.12514	0.00167
Wielkopolskie	0.3255851	0.1050285	0.001935
Zachodniopomorskie	0.5556994	0.1244597	8.01E-006
exp	0.0749304	0.010165	1.69E-013
exp^2	-0.0023292	0.0002696	$<$ 2e-16
tenure	-0.0603219	0.0101201	2.51E-009
tenure_un	0.0027389	0.0066055	0.678413
yrs_edu	-0.0835757	0.0122588	9.26E-012
female	-0.1847443	0.048633	0.000145
$\ln employ$	-0.4101964	0.053207	1.26E-014
$\ln \theta$	1.5362151	0.168068	$<$ 2e-16

Figure 4: Multiplying factor for job-termination intensity comparing the 1997 country mean.



Source: Own computation

Figure 5: Comparison of the predictions of different count models.



Source: Own computation

5 Conclusions

The conducted study provides a few interesting facts about the Polish labor market. First, there exists depreciation of the human capital, which is manifested by the predominant role of the 'up-to-date' experience in the last 10 years. The overall experience does not decrease the job-termination probability but even increases it. Such a state of affairs may be caused by the character of the Polish economy. The economic transition in the 1990's resulted in the deep change of the labor market and in consequence the experience gained before 1989 might have become less useful for the market economy reality. The negative resultant impact of employment duration in the last 10 years confirms this conjecture.

Furthermore, the analysis proved that there exists a firm-specific part of human capital which cannot be transferred across employers. The tenure of an employed individual decreases the involuntary job-separation probability, whereas the tenure at the last workplace for an unemployed person does not exhibit a statistically significant impact.

To sum up, the conducted analysis may be considered the first step in the calibration of a job-search model with heterogeneous agents for the Polish economy. Ultimately, the estimation of the remaining parameters of the fundamental reservation wage equation would enable to define reservation wages depending on the characteristics of individuals.

References

Becker, G. S. (1962): "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70.

Börsch-Supan, A. (1990): "Education and its double-edged impact on mobility," *Economics of Education Review*, 9(1), 39–53.

Burdett, K., and D. T. Mortensen (1980): "Search, Layoffs, and Labor Market Equilibrium," *Journal of Political Economy*, 88(4), 652–72.

——— (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2), 257–73.

Cameron, A. C., and P. K. Trivedi (1998): *Regression Analysis of Count Data*. Cambridge University Press.

——— (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press.

Cichomski, B., T. Jerzy, and M. Zielinski (2009): "Polish General Social Surveys: machine readable data file 1992-2008," Institute for Social Studies, University of Warsaw.

Eckstein, Z., and G. J. van den Berg (2007): "Empirical labor search: A survey," *Journal of Econometrics*, 136(2), 531–564.

Flinn, C. J. (2002): "Labour Market Structure and Inequality: A Comparison of Italy and the U.S," *Review of Economic Studies*, 69(3), 611–45.

French, E., and C. Tabe (2010): "Identification of Models of Labor Market," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4. Elsevier, 1 edn.

Greene, W. (1994): "Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models," Working Papers 94-10, New York University, Leonard N. Stern School of Business, Department of Economics.

——— (2005): "Functional Form and Heterogeneity in Models for Count Data," *Foundations and Trends in Econometrics*, 1(2), 113.

——— (2008): *Palgrave Handbook of Econometrics*vol. 2, chap. Discrete Choice Modeling. Palgrave Macmillan.

Jackman, S., C. Kleiber, and A. Zeileis (2008): "Regression Models for Count Data in R," *Journal of Statistical Software*, 27(8).

Lambert, D. (1992): "Zero-inflated Poisson regression, with an application to defects in manufacturing," *Technometrics*, 34, 1–14.

Long, S. J. (1997): *Regression Models for Categorical and Limited Dependent Variables*. SAGE, Thousand Oaks.

Mincer, J. (1962): "On-the-Job Training: Costs, Returns, and Some Implications," *Journal of Political Economy*, pp. 50–79.

Topel, R. H., and M. P. Ward (1992): "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 107(2), 439–79.

Venables, W. N., and B. D. Ripley (2002): *Modern Applied Statistics with S*. Springer, New York, fourth edn., ISBN 0-387-95457-0.

Vuong, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses," *Econometrica*, 57(2), 307–33.

Winkelmann, R. (2003): *Econometric analysis of count data, 5th Edition*. Springer.

Winkelmann, R., and S. Boes (2006): *Analysis of Microdata*. Springer.

Winkelmann, R., and K. F. Zimmermann (1998): "Is Job Stability Declining in Germany? Evidences from Count Data Models," IZA Discussion Papers 01, Institute for the Study of Labor (IZA).