CROSS-SECTIONAL ANALYSIS OF MODIFIED PHILLIPS CURVE IN THE CZECH REPUBLIC[1]

JANA HANČLOVÁ, LUBOR TVRDÝ
Department of Mathematical Methods in Economics
Faculty of Economics, VSB Technical University
Sokolská 33, 701 21 Ostrava
CZECH REPUBLIC
e-mail: jana.hanclova@vsb.cz
lubor.tvrdy@vsb.cz

Abstract

The purpose of this paper is to provide an additional empirical information about a cross-sectional analysis of the modified Phillips curve in all regions in the Czech Republic over the period 1992 - 1996. Modified Phillips curve explains the influence of the registered unemployment rate on the level of an average wage cross-sectionally. Such a dependence is investigated as a linear relationship including a modification with a weighted variable which expresses the significance of the particular districts. Further attention is devoted to a non-linear transformation based on a reciprocal or logarithmic function of the unemployment rate. We attempted to eliminate the problems occurring during the verification of the previous models by an application of a piecewise linear function. The final part of the paper deals with a cross-sectional analysis with the aim to explain the inspected dependencies more precisely.

Keywords: Modified Phillips curve, wages, unemployment rate, regional labour market model, cross-sectional analysis, sector analysis

1. Introduction

The paper deals with a dependence between a wage level and unemployment. The economic theory describes a wage inflation and unemployment like a classic Phillips curve and its modification. The main aim of the paper is not a discussion towards a theoretical problems of the curve but to confirm empirically a behaviour of the average wages in their dependency on the unemployment rate during the transformation of the Czech economy.

In the Czech Republic time series are not long enough for the time-series analysis where the values are comparable methodologically. Due to this fact we decided for a cross-sectional analysis over the sample units (i.e. regions) in a particular time period (during years). The regions are presented like districts[2] in this empirically study and we also discuss some trends of estimated coefficients during individual years 1992 -1996.

The fully comparable data without a season impacts and relatively large sample of 76 districts are the advantages of this approach from the methodological point of view. The great disadvantage of the cross-sectional analysis is an assumption of the isolation of the individual regions resp. districts and no respect with regard to their mutual

[1] This research was supported by the Grant No. 402/96/0823 from the Grant Agency of the Czech Republic

[2] New district Jeseník is included in district Šumperk.

relations that exist towards big cities (Prague, Ostrava). They are centres not only of production but consumption and this is reflected in a high mobility of the production factors with to the neighbouring regions - concerning labour force mainly. Despite above mentioned disadvantages (position of region in a space structure, the differences in an internal structure of labour force in the region) that can decrease the investigated statistical dependence we presume the application of the cross-sectional analysis at the labour market to be rational comparing to the time-series analysis.

2. Description of variables and data sources

Basic variables used in our empirical study are:

W - annual average of the nominal monthly wage (CZK) calculated by workplace-method. It means that organisations are not ranked to the district according to the location of headquarters but according to the individual business units.
Data source: Czech Statistical Office - report of Labour 3-01.

u - average annual rate of registered unemployment (%).
It holds relationship:

$$u = U / L$$

where U - arithmetic average of job applicants in individual month
L - arithmetic average of size of labour force in individual month
Data source: Ministry of Labour and Social Affairs - report of job applicants and vacancies.

3. Empirical models

3.1 Average wage depends on the unemployment rate linearly (models 1 and 2)

We regress wage on unemployment rate to find the relationship between two. Using our data samples we obtain the following results by ordinary least squares method - OLS (see table 1). Only the first model for 1992 is statistically significant at 5% level of significance but coefficient of determination (R^2) is extremely low. Estimated parameter β correspond to the economic theory. We can also reject the hypothesis that there is a serial correlation in the residuals by *Durbin -Watson statistics (DW)* at the 5% level of significance. Using our cross-series of data we also examine the validity of an assumption that the disturbances are homoscedastic (they all have the same variance). We obviously apply two formal techniques for detecting heteroscedasticity by a *Sperman's rank correlation test* (by r_s statistic) and a *Golfeld-Quandt (GQ) test*. There is no evidence of a systematic relationship between the unemployment rate and the absolute value of the residuals which might suggest that there is no heteroscedastic for our model during year 1992 (*Sig*. $r_s > 0.05$). GQ test is applicable if we assume that the heteroscedastic variance is positively related to the unemployment rate in the regression models. Since the estimated $F_{(28,28)}$ statistics (λ) do not find the critical area[3], we conclude that there is homoscedasticity in the error variance for model 1.

[3] Acceptance region F value for 28 numerator and 28 denominator df and for two-sided test at the 5% level of significant is in interval <0.4694; 2.1299>.

$$W=\alpha+\beta*u$$

Year	R^2	*Sig. level*	*DW-stat.*	α	β	r_s	*Sig.* r_s	λ
1992	0.053	**0.046**	**2.228**	4552	-52.6	-0.211	**0.068**	**1.505**
1993	0.046	0.062	**2.309**	5748	-59.0	-0.063	**0.588**	**0.852**
1994	0.025	0.174	**2.317**	6670	-47.1	-0.013	**0.910**	**0.638**
1995	0.001	0.743	**2.310**	7707	-14.1	-0.034	**0.768**	**1.151**
1996	0.001	0.753	**2.289**	9096	-15.6	-0.088	**0.450**	**1.296**

TABLE. 1: Basic estimated characteristics of the regression model 1 (OLS method).

During next stage we included weights to the model 1**.** Weighted variable s_k expresses the significance of individual regions resp. a share of workers to whom above mentioned relationship may concern. It is defined as follows:

s_k = workers in *k*th district / workers in a whole national economy.

Number of workers is valid till 31 December of the current year and with regard to the greater objectivity we have used an arithmetic average of two following years. The final regression models 2 (see table 2) are statistically significant in individual years 1992 - 96. R^2 statistics have been increased and estimated coefficients α grow up during years 1992 - 96 which is conformable to a nominal wage growth (we cannot use real wages in models as price indexes are not followed in the particular districts). Estimated β coefficients are decreasing and it represents a stronger dependence. The tests of heteroscedasticity by a *Sperman's rank correlation test* and a *Golfeld-Quandt test* conclude different results. Sperman's test cannot reject any evidence of a systematic relationship between ***u*** and absolute value of the residuals for all followed years but *GQ* test rejects heteroscedasticity hypothesis for period 1992 - 1994. Since the estimated DW value lies near 2, we can reject the hypothesis of non-zero first-order autocorrelation in the residuals at 5% level of significance.

$$W=\alpha+\beta*u \text{ - weighted by } s_k$$

Year	R^2	*Sig. level*	*DW-stat.*	α	β	r_s	*Sig.* r_s	λ
1992	0.236	**0.000**	**2.085**	4954	-130.0	-0.468	0.000	**1.996**
1993	0.251	**0.000**	**2.128**	6387	-185.2	-0.390	0.000	**1.588**
1994	0.241	**0.000**	**2.089**	7591	-209.6	-0.374	0.001	**1.649**
1995	0.195	**0.000**	**2.041**	8945	-253.0	-0.346	0.002	3.171
1996	0.210	**0.000**	**2.037**	10671	-312.9	-0.374	0.001	3.210

TABLE. 2: Basic estimated characteristics of the regression model 2 (WLS method).

Further more, we tried to change a functional form through non-linear transformations which should consider more rigid wages for higher ***u***.

3.2 Non-linear functional form of influence u on W (models 3 and 4)

In the first part we will pay attention to reciprocal *1/**u*** non-linear function in a model 3 (see table 3).

$$W = \alpha + \beta / u$$

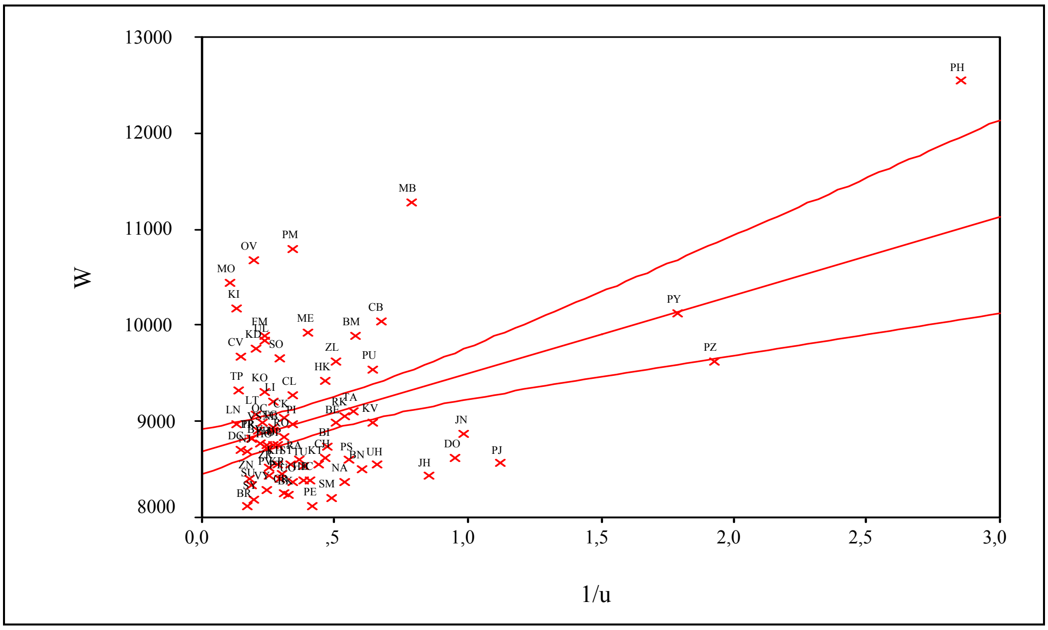
Year	R^2	Sig. level	DW-stat.	α	β	r_s	Sig. r_s	λ
1992	0.068	**0.023**	**2.316**	4256	299.2	-0.170	**0.141**	**0.709**
1993	0.125	**0.002**	**2.141**	5392	362.9	-0.057	**0.627**	**1.174**
1994	0.185	**0.000**	**2.256**	6269	551.0	-0.069	**0.556**	**1.568**
1995	0.178	**0.000**	1.518	7395	557.2	-0.135	**0.245**	**0.869**
1996	0.192	**0.000**	**1.832**	8680	816.3	-0.175	**0.130**	**0.771**

TABLE. 3: Basic estimated characteristics of the regression model 3 (OLS method).

The estimated coefficients α, β and statistical characteristics R^2, *Sign. level* for particular regression models 3 are presented in a table 3. The regression statistics look satisfactory at the first glance in years 1992 - 96.

Visually (see graph 1), without more tests, we can see that regression include strong influence of outliers. We measure the change in estimates of the regression coefficients in consequence of removing one case by Cook's distance[4]. The main outlier is Prague (PH) with Cook's distance = 3,31 but Cook's distance median is only 0,0034 in comparison with the share of Prague district on labour force in all national economy which is 15,7% (i.e. s_{PH}) but median value is 0,98% in 1996.

We also cannot reject hypothesis of zero autocorrelation at 5% level of significance except 1995 year. For our reciprocal regression models the *Sperman's* and also *GQ tests* conclude the same results for detecting heteroscedasticity. We cannot reject hypothesis that the error variance does not depend on level of unemployment rate.



GRAPH 1: *W* depends on *1/u* for year 1996 including estimated regression function and confidence interval for mean *W*.

Further non-linear function to prevent the problem of outliers (for values of unemployment rate lower 1 mainly) is the natural logarithmic function. It is assumed for this model 4 that estimator $\beta < 0$ and for estimated regression $W > 0$. The results of this model do not show outstanding improvements (see table 4). We identified weak

[4] Cook can be viewed as a scaled Euclidean distance measure between two vectors of fitted values in the first, the district is included; in the second, the district is excluded (see p.180, SPSS, 1997).

dependence ***W*** on *ln(**u**)*. Level of estimated intercepts α increases and negative value of the estimated slope coefficients β decreases in time (with the exception of 1995).

We can reject hypothesis of zero first-order autocorrelation. Due to the outliers problem the influence of Prague is weaker in comparison to the previous reciprocal model but still is strong enough (Cook's distance for Prague is 1,5633 in 1996 comparing to median value that equals 0,0038).

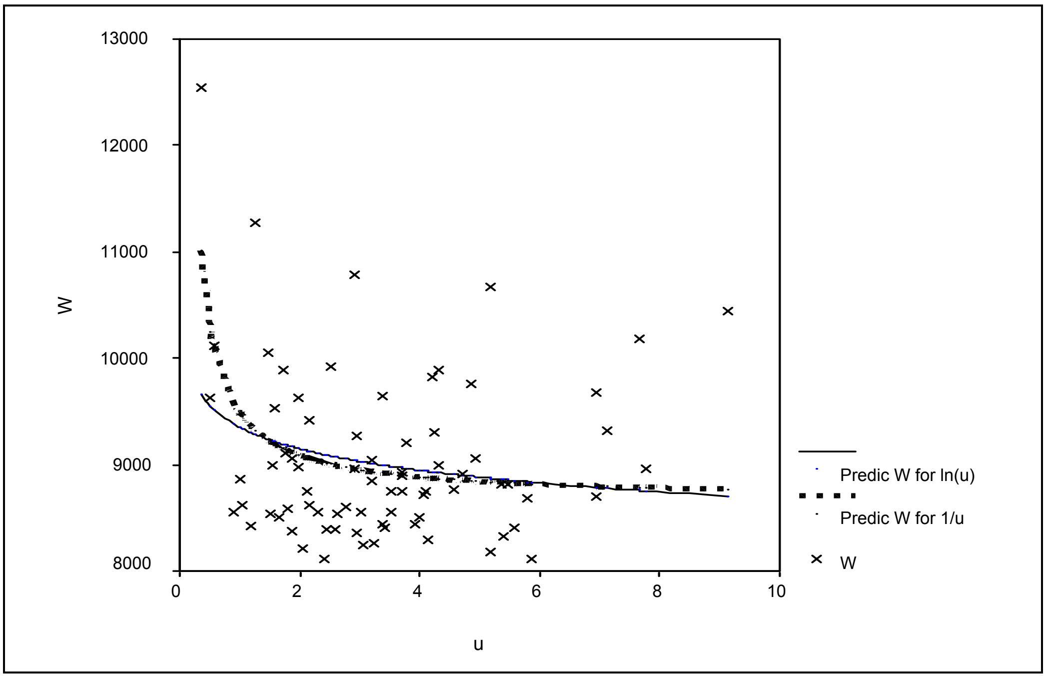
There is no evidence of a systematic relationship between the unemployment rate and the absolute value of the residuals which might suggest that there is no heteroscedastic problem for our model 4 during years 1993 - 96. For the *GQ test* the estimated $F_{(28,28)}$ statistics do not find the critical area. We can conclude that there is homoscedasticity
in the error variance for followed model 4 during the same years.

$$W=\alpha+\beta*ln(u)$$

Year	R^2	*Sig. level*	*DW-stat.*	α	β	r_s	*Sig.* r_s	λ
1992	0.057	**0.038**	**2,206**	4545	-154.8	-0.231	0.044	**1.476**
1993	0.079	**0.014**	**2,310**	5768	-205.7	-0.082	**0.482**	**0.944**
1994	0.075	**0.017**	**2,323**	6768	-240.3	-0.087	**0.454**	**0.776**
1995	0.045	0.065	**2,344**	7874	-211.8	-0.088	**0.452**	**1.415**
1996	0.055	**0.041**	**2,316**	9350	-291.1	-0.178	**0.124**	**1.526**

TABLE 4: Basic estimated characteristics of the regression model 4 (OLS method).

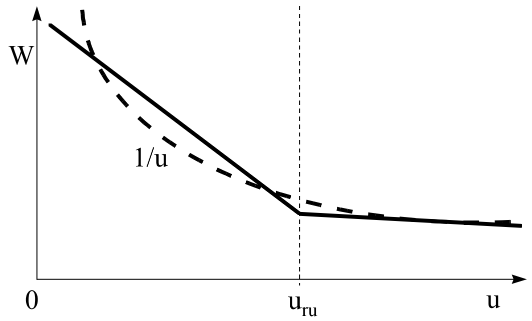
It is necessary to submit that neither logarithmic function did not prevent the problem of outliers. It was caused by data non homogenity probably. We present graphs (see graph 2) to compare estimated non-linear curves of the models 3 and 4 for year1996.



GRAPH 2: The relationship between *W* and *u* with estimated regression functions by modles3 and 4 for 1996.

3.3 The form of dependence W on u piecewise linear function (model 5)

The other possibility of the modification of the previous models is an approximation of the non-linear relation by piecewise linear function (see graph 3). Another problem arises about setting the limits of break of line u_{ru} for estimation of the individual lines. It is determinated by the range of the selected sample. In the following model 5 we estimate regression lines for the different limits u_{ru}, where ***ru*** indicates the order of ***u*** according to the size (ranked upward for the districts of interest) ***ru*** equals 30 means the lowest limit suitable for statistical verification.



GRAPH 3: Approximation of the non-linear model by piecewise linear model.

$$W=\alpha+\beta*u$$

$$\text{for}$$

$$u\in\langle 0;u_{ru}\rangle \text{ where } ru\in\{30;35;40;45;50;55;60\}$$

	ru=30	ru=35	ru=40	ru=45	ru=50	ru=55	ru=60
R^2							
1994	**0.252**	**0.119**		**0.083**[5]	**0.102**	**0.123**	**0.088**
1995	0.176	0.2206	0.1386		0.084	0.099	0.099
1996	**0.258**	**0.136**	**0.169**	**0.164**	**0.148**	**0.130**	**0.083**
β							
1994	-432	-263		-179	-170	-163	-126
1995	-507	-490	-354		-210	-202	-179
1996	-783	-488	-472	-407	-339	-282	-201
α							
1994	7331	7088		6961	6944	6931	6856
1995	8550	8528	8349		8131	8117	8075
1996	10470	10060	10035	9933	9816	9711	9555
r_s							
1994	0.654	0.308		**-0.037**	**-0.076**	**-0.108**	**-0.187**
1995	0.610	0.596	0.365		**0.010**	**0.006**	**-0.036**
1996	0.714	0.440	0.420	0.317	**0.159**	**0.035**	**-0.085**

TABLE 5: Basic statistical characteristics for model 5 in dependence u_{ru} value (by OLS method).

[5] Level of significance is 5,5%.
[6] Durbin -Watson statistics are in zone of indecision.

Table 5 shows statistical characteristics of estimated regression functions that have been verified statistically significant at the 5% level of significance (for an exception see note no. 5). At the same time the extra bold $\boldsymbol{R^2}$ values rejected statistically significant first order autocorrelation of residuals by *DW statistic*. We detect heteroscedasticity by a *Sperman's rank correlation test* (by $\boldsymbol{r_s}$ statistic). There is no evidence of systematic relationship between the unemployment rate and the absolute value of the residuals which might suggest that there is no heteroscedasticty for our piecewise linear model 5 for break point ***ru*** equals 50, 55 and 60.

Table 5 also demonstrates that the values of estimated β is decreasing, when the break point ***ru*** increases. It means that dependence of wage on rate of unemployment falls down. The influence of Prague outliers is still important but it is comparable to its weight $\boldsymbol{s_k}$ (for break point ***ru*** *= 30).*

The estimated regression line for $\boldsymbol{u} \in (\boldsymbol{u_{ru}};\infty)$ (the area with a low dependence of ***W*** on ***u***) was not statistically significant mostly as it could be expected (see graph 3). Because of the low levels of estimated β coefficients statistical significance for ***ru*** > 45 is lower than 5 % level in 1995 only ($\boldsymbol{R^2}$ = 0,166; α = 6608 a β = **197**). This estimated β coefficient is positive that is out of keeping with the economic theory. We can explain it by the existence of the districts where higher wage is paid than average wage level and higher unemployment rate exists (for example districts Karviná, Most, Ostrava, Frýdek-Místek, Kladno, Teplice, Chomutov). This higher wage level is determinated by a higher share of employment concentrated in heavy or mining industry with relatively high wages.

3.4 Cross-sectional sector analysis (model 6)

Based on the empirical results we can submit that very weak dependence of wages on registered unemployment exists. One cause could be labour market segmentation (see Sirovátka, 1997). We could distinguish two segments - primary and secondary (marginally). The primary segment includes workers with higher professional education or the other specific ability. The secondary segment is created by unskilled workers which are interchangeable easily in the different professions. Therefore there is a higher pressure of unemployeed people on the level of wage in the secondary segment. We assume that the relative segments rate is different in particular sectors. Therefore we expect different intensity of linear dependencies of sectored wages on district unemployment rate. We chose the following sectors for our empirical tests:

a) agriculture, fishing, forestry;
b) industry (processing industry, power supply and mining industry);
c) building industry;
d) retails, trades, car mending;
e) transport, store posts, telecommunication;
f) banking and insurance branch;
g) education sector;

We assume the negative linear dependence for the first five above mentioned sectors.

Model parameters presented in table 6 are statistically significant at the 5% level of significance. Extra bold values of $\boldsymbol{R^2}$ show where the hypothesis about zero first order autocorrelation of residuals was not rejected. *GQ test* rejects heteroscedasticity hypothesis for extra bold values of λ.

$$W_{sector}=\alpha+\beta*u$$

R^2	agriculture	building industry	retails, trades	transport, store posts, telecom.	banking, insurance branch
1992		**0.174**	0.214	**0.119**	
1993		**0.284**	**0.269**	**0.067**	
1994	0.096[7]	0.254[7]	**0.277**	**0.087**	
1995	**0.173**	0.124	**0.128**	0.099	0.057
1996	**0.208**	0.074	0.111		
α					
1992		5115	4087	4538	
1993		6800	5304	5519	
1994	6073	7849	6412	6738	
1995	7312	8881	6922	7924	12455
1996	8408	9972	8340		
β					
1992		-109	-112	-58	
1993		-179	-180	-45	
1994	-65	-168	-225	-101	
1995	-117	-152	-193	-92	-159
1996	-157	-115	-222		
λ					
1992		**0.644**	2.697	**1.177**	
1993		**1.384**	4.099	2.263	
1994	**0.588**	**0.861**	3.541	5.026	
1995	**1.053**	2.207	2.249	3.611	9.941
1996	**1.396**	**1.595**	**1.821**	5.255	

TABLE 5: Basic statistical characteristics for model 6 for selected (by OLS method).

We can observe the growing dependence in agriculture since 1994. On contrary to our expectations we did not prove the dependence statistically in the industry. Building industry shows the dependence during 1992 - 96 with the top in 1993. Relatively high $\boldsymbol{R^2}$ values were observed in the trade sector with the top in 1994. The estimated β coefficients fluctuate during time and problem with the residual autocorrelation occurs. The weak dependence is in transport. The relationship has not been proved in banking industry (excluding 1995) and in sector of education but we presumed it.

4. Conclusion

The aim of this paper was to investigate empirically the dependence of wages on registered unemployment rate. We used cross-sectional analysis in the Czech Republic during 1992 -1996.

Our investigation was based on simplified linear dependence which did not prove statistically significant relationship. We applied weighted variable and gain a strong improvement of the estimated linear model. At the same time the problems with

[7] Durbin -Watson statistics is in interval of indecision.

the heteroscedastic residuals appeared. We changed the form of regression function due to the lack of suitable explanatory variables. Non-linear transformation based on reciprocal or logarithmic regression function caused the problem of outliers. Then we introduced a piecewise linear function. We devoted our attention to the lower levels of unemployment rates when wages are less rigid usually and we summarised relatively satisfactory results for the period 1994 -1996.

The final part of the paper deals with the dependence of unemployment on wages in selected sectors according to the theory of labour market segmentation. We found out the positive trend of dependence between wages and unemployment rate in the region in agriculture mainly. It can mean the start of market relations in the agriculture sector in the Czech economy. The wage level in building industry has been determinated by the unemployment rate, too. We can presume that wages in industry are also determinated by the other factors not only by the unemployment rate.

References:

Čihák, M.: Strukturální Phillipsova křivka. Politická ekonomie, 1996, no. 2, p.240-247.

ČSU: Okresy České republiky v roce 1994, 1995, 1996. Praha, ČSU 1995, 1996,1997.

ČSU: Pracovníci a mzdové fondy v krajích a okresech podle odvětví národního hospodářství v České republice za rok 1992. Praha, ČSU 1993.

ČSU: Zaměstnanost a mzdy pracovníků v evidenčním počtu organizací v krajích a okresech ČR za rok 1994. Praha, ČSU 1995.

Gujarati, D. N.: Basic Econometrics. McGraw-Hill, Inc. 1995.

MPSV: Uchazeči a volná pracovní místa. MPSV for yers 1992-1996.

Sirovátka, T.: Segmentační tendence a procesy na trhu práce v České republice. Politická ekonomie, 1997, no. 6, p. 804-816.

SPSS: SPSS Base 7.5 Applications Guide. Chicago, SPSS 1997.