Bank i Kredyt 44 (2), *2013, 159–174*

BANK I KREDYT

www.bankikredyt.nbp.pl
www.bankandcredit.nbp.pl

Regional wage determinants in Poland: the empirical verification of the NEG approach

Andrzej Cieślik*, Bartłomiej Rokicki

Submitted: 24 May 2012. Accepted: 15 October 2012.

Abstract

The paper aims at explaining the regional wage determinants in Poland using the New Economic Geography approach. We test the hypothesis that regional wages are positively related to the regional economic potential, controlling for individual worker, firm, industry and regional characteristics. We demonstrate that the regional economic potential strongly influences the regional level of wages. Moreover, we find that individual characteristics also play an important role in wage determination in Poland.

Keywords: economic potential, wages

JEL: J31, R12

* University of Warsaw, Department of Economics; e-mail: cieslik@wne.uw.edu.pl.

1. Introduction

Following the emergence of the so-called New Economic Geography (NEG) literature initiated by Krugman (1991) the uneven distribution of economic activity across space has become one of the most interesting and controversial research topics in economics in the last twenty years. In this paper, we examine the spatial distribution of wages in Poland using the NEG approach that underlines the importance of a region's proximity to markets in other regions in the determination of regional wages. In particular, we test the research hypothesis stating that regional wages are positively related to the regional economic potential, having controlled for individual worker characteristics, firm, industry and regional characteristics.

The main idea that the level of economic activity in a region depends on its access to markets for produced goods, is widespread in the theoretical NEG literature.[1] In particular, Krugman (1991; 1992) explains agglomeration formation on the basis of the interaction between transport costs and increasing returns to scale at the level of firms. In his models, demand linkages between regions are created by scale economies and transport costs. These linkages contribute to spatial agglomeration which results in greater production and higher equilibrium rewards in locations close to large markets.

The Krugman (1992) model is closely related to Harris (1954) economic potential function, which did not have microeconomic foundations but was based on the intuitive idea that the demand for goods produced in a location is the sum of the purchasing power in all other locations, weighted by transport costs. Fujita, Krugman and Venables (1999) reformulated the economic-potential concept by showing how it can be derived from a variety of formal spatial models. The market potential function originally derived from the Krugman model and extended to many regions indicates that wages in a region are increasing in the income of surrounding regions, decreasing in transportation costs to these regions, and increasing in the price of competing tradable goods in these regions.

However, the main disadvantage of the approach proposed by Krugman (1991) and his followers was the fact that it limited the concentration of workers in one region only with the transport costs and it did not take into account other factors that could be the result of such an increase in the number of inhabitants in a given region. This issue was taken up by Helpman (1998) who created a model in which agglomeration increases the demand for local services that are non--tradable across regions, such as housing. As a result, locally supplied housing services generate a major dispersion force. A decrease in the housing supply leads to an increase in house prices and declining the real wages of workers in the region.

Hanson (2005) used the NEG model developed by Helpman (1998) and derived the market potential functions. Different from the market potential function of Fujita, Krugman and Venables (1999), Hanson model (2005) controls for the regional variation in the price of non-traded goods. Wages in his model are increasing in the income of surrounding regions and the supply of housing stock and decreasing in transportation costs to these regions. Also, higher wages in nearby regions will also increase the wage in the surrounded region: higher wages in surrounding regions raise the relative price of traded goods produced in these regions, which increases their demand for goods produced in the surrounded region. Higher production further raises the region's demand

[1] For example, see Fujita, Thisse (1996); Ottaviano, Puga (1998); Fujita, Krugman, Venables (1999); Brakman, Garretsen, Marrewijk (2001); Baldwin et al. (2003) and Combes, Mayer, Thisse (2008) for the survey of the earlier literature.

for labor and its nominal wages and housing prices. Therefore, the Hanson-Helpman model offers a meaningful description of the spatial feature of an economy, that has been employed in many empirical studies.

Basing on the Helpman-Hanson model in this paper we examine the spatial correlation of wages and the market access across regions in Poland in 2007 using individual data that comes from the Polish labor force survey. First, we estimate the relationship between regional wages and the market potential function that is currently considered to be a reduced form for many NEG models. Then, we control for individual characteristics of workers, industries, regions and firms. We find that regional economic potential strongly influences the regional wage levels. Moreover, we show that many individual characteristics of workers and firms play a very important role in wage determination in addition to the regional economic potential. Therefore, this paper combines the fields of economic geography with that of labor economics within a single empirical framework.

The reminder of the paper is organized as follows. In section 2, we provide the review of the relevant literature on the empirical relationship between market access and wages. In section 3 we discuss the analytical framework. In section 4, we describe the dataset. In section 5, we discuss the estimation results. Section 6 summarizes and concludes.

2. Literature review

Initially, in the 1990s the NEG literature was mainly a theoretical enterprise. According to Redding (2010, p. 298) "although a rich and extensive body of theoretical research on new economic geography has now emerged, empirical research remains comparatively less developed".[2] In particular, there have been few empirical studies trying to investigate the relationship between regional wages and the economic potential of the region. However, the situation has been gradually changing since the mid-2000s but the majority of previous studies concentrate mainly on developed countries in North America and Western Europe.

In particular, Hanson (2005) examined the spatial correlation between wages and consumer purchasing power across U.S. counties to see whether regional demand linkages contribute to spatial agglomeration. He estimated both a simple market-potential function and an augmented market-potential function derived from the Krugman model of economic geography for the period 1970, 1980 and 1990 in order to explore the importance of scale economies and transport costs. His estimation results suggested that regional differences in wages are associated with the proximity to large markets. Also, his estimation results showed that the data was in favor of the augmented market-potential function derived from the theory.

Brakman, Garretsen and Schramm (2004) employed the Helpman-Hanson model to empirical study using data for Germany in 1995. Their estimation yielded results that also verified positively the correlation between spatial wage distribution and the geographical distribution of economic activities. Their studies extended the Helpman-Hanson framework in two major ways. First, they estimated the model not only using data on the housing stock but also land prices. They argued

[2] See Head, Mayer (2004); Overman, Redding, Venables (2003); Surico (2003) and Redding, Venables (2004) for the early surveys on the economic geography literature. A more recent survey on the empirics of the new economic geography literature has been provided by Redding (2010).

that the use of land prices is more consistent with the model, as the non-tradable good is set as the dispersion force in the model. Second, they dropped the assumption of real wage equalization. They found clear-cut evidence for both the spatial wage structure and the relevance of structural parameters of the model.

The research on Germany was done also by Roos (2006) who used a market potential function to examine the spatial correlation of wages and consumer purchasing power across regions in West Germany and found that wages of skilled workers showed positive correlation with purchasing power in other regions. Further support for the empirical relationship between wages and market access using data on regions of the European Union was provided by Breinlich (2006) and Head and Mayer (2006).

The empirical evidence on the relationship between the spatial wage structure and economic potential for the new members of the European Union that joined the EU in 2004 and 2007 during the two waves of the Eastern Enlargement until now remains scarce. Notable exceptions include Brülhart and Koenig (2006) who analyzed the internal spatial wage and employment structures of the Czech Republic, Hungary, Poland, Slovakia and Slovenia, using regional data for 1996–2000 and Damijan and Kostevc (2011) who examined the adjustment pattern of relative regional wages in Bulgaria, Estonia, Hungary, Romania and Slovenia in the period 1990−2004. However, following the previous literature for the developed Western economies these authors focused on only on determination of average wages and did not control for worker and firm characteristics.

Moreover, although there seems to be clear evidence of a positive empirical link between wages and market access a key challenge for the literature has been to establish that this relationship is causal. According to Redding (2010, p. 301) "… it is difficult empirically to disentangle the effects of market access from other leading determinants of comparative economic development such as locational fundamentals or institutions". Further, he argues that in order to empirically disentangle market access from these other determinants one requires exogenous variation along at least one dimension. Therefore, one possible approach is to use instruments for market access, such as lagged population levels or growth rates.

Therefore, this study attempts to close an existing gap in the empirical NEG literature in two important ways. First, in particular, in contrast to the previous studies that use average regional wages in our paper we use individual wage data that allows us controlling for worker and firm characteristics. Second, we endogenize the concept of economic potential of the region using variables postulated by the Helpman-Hanson model in our estimating equations.

3. Analytical framework

In our study we endogenize the value of the economic potential and estimate the economic potential of the region as the function of the regional housing stock, income and employment postulated by the Helpman-Hanson model. Hence, we simultaneously estimate a system of two equations. The first of them, expressed in logs takes the form of:

$$\ln P_i = \theta \ln I_i + \omega \ln E_i + \rho \ln H_i + \varepsilon \tag{1}$$

where:

P_i – the economic potential of location i,
I_i – regional income,
E_i – regional employment,
H_i – housing stock.

In our study the economic potential P_i for location i is calculated as a sum of a measure of economic activity in location i and a weighted sum of a measure of economic activity of all other locations j with the weights declining with distance:

$$P_i = M_i + \sum_{j=1}^{n} w_{ij} M_j \tag{2}$$

where:

$M_{i,j}$ – a measure of the volume of economic activity in location i and j,
$w_{i,j}$ – spatial weights.

The second term in the above equation is usually referred to as a spatial lag in M_i calculated using the inverse-distance minmax-normalized spatial-weighting matrix.[3] This matrix was afterwards applied to compute the weighted averages of variable M.

The second estimating equation expressed in logs is:

$$\ln \varpi_{ij} = \alpha \ln P_i + \beta \ln X_j + \gamma \ln Y_j + \mu_i + \varepsilon \tag{3}$$

where:

$\varpi_{i,j}$ – the nominal wage rate of an individual j in location i,
P_i – the economic potential of location i,
μ_i – a location-specific individual effect,
X_j – individual characteristics (gender, age, education, etc.),
Y_j – non-individual characteristics (ownership, section, firm size, etc.),
ε – a white noise disturbance term.

Our system of estimations is estimated using the standard two-stage least squares estimation method.

4. Data sources

The data for our paper are drawn from a variety of sources. This section describes the sources and the manner in which we combine these data in our empirical study, and provides a discussion of some descriptive statistics.

[3] See Kelejian, Prucha (2010) for an introduction to normalization methods.

4.1. Wages

For this analysis, a cross-section micro-data set has been created covering almost 10 thousand individual observations in the fourth quarter of 2007. All individual data (including wages) comes from the Polish Labour Force Survey.[4] The highest wages are found in the (capital-city) Mazowiecki voivodeship that is the best developed region in Poland. It is followed by Śląski, Dolnośląski and Małopolski voivodeships that also may be considered as well developed. The spatial distribution of wages in our dataset is quite similar to the distribution of average gross wages in 2007.[5] This is shown in the Table 1 where regional distribution of net wages in our data set (after weighting) and gross wages from the Central Statistical Office (CSO) data are compared.

The average distribution of net wages in our dataset is far from being normal, hence for estimations we use log of wages that has a distribution close to normal. Below we show the histogram of log of wages in our sample.

It is important to note that recent advances in the spatial econometrics indicate the necessity for using such techniques in the case of the existence of spatial correlation. Before applying formal statistical tests we plot the spatial distribution of net wages on the map of Poland. It does not necessarily indicate the existence of spatial correlation.

In order to formally test for the existence of spatial correlation we computed the Moran's I and Geary's c statistics for net wages. As shown in Table 2, the performed tests do not unambiguously confirm the existence of spatial correlation. Hence, we reject the null hypothesis of spatial correlation of net wages and estimate the empirical model using the traditional instrumental variables approach.

4.2. Economic potential

It has been frequently argued in the literature that transport cost is preferred to geographic distance. Unfortunately, calculating transport costs always involves several arbitrary assumptions concerning appropriate cost measures, modes of transport and types of goods. For example, Clark, Wilson and Bradley (1969) assumed a 200-mile limit for road transport, with journey over this distance being handled by rail, while Keeble, Owens and Thompson (1982) concentrated entirely on road transport and simply computed shortest distances between regional and national modes. In order to avoid such arbitrary assumptions, we measure distance between regions using the geospatial data based on the geographical coordinates describing the boundaries of the spatial units. This data comes from the ESRI shapefile[6] for Poland.

Calculating the value of economic potential requires us to specify our measures of economic activity and distance. Many previous economic potential studies relied on values of regional gross domestic product (GDP). However, economic geography studies used also population, income or retail sales. Harris (1954) noted that for international comparisons national income serves well. In the present research we use regional GDP expressed at market prices and measured in PLN.

[4] The LFS includes more than 40 thousand observations, yet less than 10 thousand contains data concerning wages.
[5] The data on gross wages is systematically published by the Central Statistical Office. The important methodological difference is that the data comes from forms filled in by employers.
[6] The ESRI shapefile is a popular geospatial vector data format for geographic information systems software.

This data comes from the annual database compiled by the Central Statistical Office available online.

The economic potential variable in our two-stage regression is instrumented with the use of three variables predicted by the Helpman-Hanson model: the gross regional disposable income, the regional stock of housing and the regional employment. The regional disposable income is expressed at current market prices in millions of PLN. The regional housing stock per capita is measured in squared meters per person. Finally, the regional employment is measured as the number of people employed. All instrumental variables come from the annual database complied by the Central Statistical Office available online.

4.3. Control variables

All the control variables measuring individual characteristics come from the Polish Labour Force Survey. The control variables include: gender, age, age squared, experience, experience squared, education, marital status, supervisory functions, agglomeration economies, firm size, firm ownership, industry and regional specific effects. The gender is measured using an indicator variable that takes value 1 for men and 0 for women. Age and age squared are expressed in years. Similarly, experience and experience squared are expressed in years. The education variable that comprises seven categories ranging from the highest to the lowest: Ph.D., tertiary (both BA and MA), postsecondary, higher vocational, higher general, vocational, junior higher (*gimnazjum*), primary, and incomplete primary. The marital status includes four categories: single, married, widowed, divorced. The benchmark category is single. The supervisory function variable takes two values: 1 if a person is performing a supervisory function over work of other people and 2 otherwise. Agglomeration economies are proxied using the variable measuring the size of the living place. This variable includes eight categories ranging from the largest city with over 100,000 inhabitants to a village. The firm size is measured in terms of the firm employment using variables that include sixteen categories ranging from the self-employed to a firm with more than 101 employees. The benchmark category is self-employed. The firm ownership is measured using an indicator variable that takes value 1 if a firm is privately owned, and 0 if state owned. Finally, we also use indicator variables to control for industry and regional specific effects.

5. Estimation results

As mentioned before, the previous empirical studies, such as Hanson (2005) or Roos (2006), analyzed the relationship between the spatial wage structure and economic potential using average regional wages. In this paper we use individual wage data that allows us controlling for both worker and firm characteristics. In Table 3 we report the results of estimation that are summarized in five different specifications. The estimations were made using the two-stage least squares method. In Table 3 we report the estimation results obtained from the second-stage only.[7]

[7] The first-stage estimation results are shown in the Table 4. All the explanatory variables used in the first-stage regression were statistically significant at the 1% level and showed the expected signs that were in line with the predictions of the Helpman-Hanson model. These variables explained 93% of the variation in the economic potential variable.

In column (1) we report estimation results from the benchmark 'pure NEG approach' specification in which regional wages are related only to the regional economic potential. The estimated coefficient on the economic potential variable is positive and statistically significant already at the 1% level which is consistent with the predictions of the NEG theory. This result confirms the significant role of the economic potential in the wage determination in Poland and justifies the use of the NEG approach. However, the empirical model with the regional economic potential as the sole explanatory variable is able to explain only about 22% of wage variation. Therefore, in subsequent columns of Table 3 the benchmark specification is extended to include additional explanatory variables postulated by the labor economic literature.

The estimation results obtained from the specification of the model that includes basic characteristics of an individual worker such as gender, age, educational attainment and marital status are reported in column (2). These results are in line with the standard predictions of labor economics. In particular, the estimation results show that the gender premium exists, i.e. women earn less the men, the wage is increasing with age and educational attainment and that married workers earn more than single, widowed or divorced. All the aforementioned additional explanatory variables are statistically significant at the 1% level and their inclusion increases the explanatory power of the empirical model to over 45%. Moreover, the estimated coefficient on the economic potential variable is also statistically significant at the 1% level and remains of a similar magnitude as in the benchmark specification in column (1).

In column (3) we report estimation results obtained from the specification the controls for individual characteristics of workers linked to their professional activities which include work experience, work experience squared and supervisory functions. The estimated coefficients on both work experience and supervisory functions are statistically significant already at the 1% level and display expected positive signs while the estimated coefficient on the work experience squared variable is not statistically significant at all. The explanatory power of the model increases to 49%.

The estimation results obtained from the extended specification of the model that controls for agglomeration externalities by adding territorial variables such as the living place and regional indicator variables are reported in column (4). All the estimated coefficients on there variables are statistically significant, although at different levels of statistical significance. These results confirm the existence of agglomeration externalities by showing that wages are positively correlated with the size of the city (the bigger city the higher wages). Moreover, the regional dummy variables are also statistically significant in the majority of cases and display negative signs. This means that for most regions negative regional 'wage premiums' exist with respect to the (capital-city) Mazowiecki voivodeship. The explanatory power of the model rises to over 57%.

Finally, in last specification reported in column (5) we controlled also for employers' characteristics. These included the economic section (at NACE 2-digit level, agriculture is a benchmark section), the economic sector (public or private ownership) and the firm size. The estimation results show the existence of firm size premiums, i.e. the bigger firm the higher wage it pays. The private sector also seems to offer higher salaries than the public sector. The surprising result may be that we did not find any industry premiums as all the estimated coefficients on the dummy variables for the particular sections of the economy were not statistically insignificant. This may be due to the fact that these sections are quite highly aggregated and the actual wage variation can exit within those sections rather between them. Another possible explanation may be

the low quality of the data that comes from the wage part of the LFS, as argued by various authors (e.g. Bukowski 2008; Newell, Socha 2007). Still, the explanatory power of the model increases to over 60%.

6. Concluding remarks

The paper aimed at explaining the determinants of regional wage disparities in Poland using the New Economic Geography approach combined with the labor market approach that allowed us controlling for a range of individual characteristics of workers as well as of employers. We tested the hypothesis postulated by the NEG literature that regional wages are positively related to the regional economic potential with and without controlling for individual worker characteristics, industry and regional characteristics. We demonstrated that the regional economic potential is positively related to the level of wages in the region and statistically significant at the 1% level across the specifications. This result is in line with the predictions of the NEG literature. Moreover, we also found that the vast majority of individual worker characteristics played an important role in wage determination which is in line with the standard labor market literature on wage determinants. It is interesting to note that the economic potential variable alone accounts for about one third of the explained wage variation and the rest is explained by the individual characteristics.

References

Baldwin R., Forslid R., Martin P., Ottaviano G.M., Nicoud F.R. (2003), *Economic geography and public policy*, Princeton University Press, Princeton.

Brakman S., Garretsen H., Marrewijk C.V. (2001), *An introduction to geographical economics*, Cambridge University Press, Cambridge.

Brakman S., Garretsen H., Marrewijk C.V. (2009), Economic geography within and between European nations: the role of market potential and density across space and time, *Journal of Regional Science*, 49(4), 777–800.

Brakman S., Garretsen H., Schramm M. (2004), The spatial distribution of wages: estimating the Helpman-Hanson model for Germany, *Journal of Regional Science*, 44(3), 437–466.

Breinlich H. (2006), The spatial income structure in the European Union – what role for economic geography?, *Journal of Economic Geography*, 6(5), 593–617.

Brülhart M., Koenig P. (2006), New economic geography meets Comecon – regional wages and industry location in Central Europe, *Economics of Transition*, 14(2), 245–267.

Bukowski M. (2008), *Zatrudnienie w Polsce 2007. Bezpieczeństwo na elastycznym rynku pracy*, Ministerstwo Pracy i Polityki Społecznej, Warszawa.

Clark C.F., Wilson F., Bradley J. (1969), Industrial location and economic potential in Western Europe, *Regional Studies*, 3, 197–212.

Combes P.-P., Mayer T., Thisse J.-F. (2008), *Economic geography. The integration of regions and nations*, Princeton University Press, Princeton.

Damijan J.P., Kostevc Č. (2011), Trade liberalization and economic geography in CEE countries: the role of FDI in the adjustment patterns of regional wages, *Post-Communist Economies*, 23(2), 163–189.

Fujita M., Krugman P., Venables A. (1999), *The spatial economy: cities, regions, and international trade*, MIT Press, Cambridge.

Fujita M., Thisse J.F. (1996), *Economics of agglomeration, Journal of the Japanese and International Economies*, 10(4), 339–378.

Hanson G. (2005), Market potential, increasing returns and geographic concentration, *Journal of International Economics*, 67(1), 1–24.

Harris C.D. (1954), The market as a factor in the localization of industry in the United States, *Annals of the Association of American Geographers*, 44(4), 315–348.

Head K., Mayer T. (2004), The empirics of agglomeration and trade, in: V. Henderson, J.-F. Thisse (eds.), *Handbook of regional and urban economics*, North-Holland, Amsterdam.

Head K., Mayer T. (2006), Regional wage and employment responses to market potential in the EU, *Regional Science and Urban Economics*, 36(5), 573–595.

Helpman E. (1998), The size of regions, in: D. Pines, E. Sadka, I. Zilcha (eds.), *Topics in public economics*, Cambridge University Press, Cambridge.

Keeble D., Owens P.L., Thompson C. (1982), Regional accessibility and economic potential in the European Community, *Regional Studies*, 16, 419–432.

Kelejian H.H., Prucha I.R. (2010), Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, *Journal of Econometrics*, 140(1), 53–130.

Krugman P. (1991), Increasing returns and economic geography, *Journal of Political Economy*, 99(3), 483–499.

Krugman P. (1992), *A dynamic spatial model*, NBER Working Papers, 4219.

Newell A., Socha M.W. (2007), The Polish wage inequality explosion, *Economics of Transition*, 15(4), 733–758.

Ottaviano G.I.P., Puga D. (1998), Agglomeration in a global economy: a survey of the 'New Economic Geography', *World Economy*, 21(6), 707–731.

Overman H.G., Redding S., Venables A. (2003), Economic geography: a survey of empirics, in: E.K. Choi, J. Harrigan (eds.), *Handbook of international trade*, Blackwell, Malden.

Redding S. (2010), The empirics of New Economic Geography, *Journal of Regional Science*, 50(1), 297–311.

Roos M. (2006), *Wages and market potential in Germany*, Discussion Papers in Economics, 00-09, University of Dortmund, Department of Economics.

Acknowledgements

The authors would like to thank the participants of the 4th GfR Summer Conference in Regional Science in Dresden for their helpful comments, suggestions and discussions on the earlier draft of this paper. Financial support of the Polish National Science Centre under the grant number 2011/03/B/HS4/05681 is gratefully acknowledged.

Appendix

Table 1
Regional distribution of net and gross wages in Poland

Voivodeships	Net wages (our set)		Gross wages (employers data)	
	Mean (std. error)	Rank	Mean	Rank
Dolnośląski	1226.72 (28.76)	3	2860.66	4
Kujawsko-pomorski	955.33 (27.50)	9	2443.17	13
Lubelski	1057.47 (22.92)	7	2470.86	11
Lubuski	910.96 (16.30)	10	2486.22	10
Łódzki	631.28 (22.53)	16	2429.97	14
Małopolski	1144.50 (38.51)	4	2666.24	5
Mazowiecki	2268.37 (59.83)	1	3670.84	1
Opolski	718.84 (24.14)	14	2607.45	8
Podkarpacki	783.82 (19.21)	11	2372.82	16
Podlaski	783.77 (23.27)	12	2524.95	9
Pomorski	1143.06 (38.08)	5	2882.56	3
Śląski	1575.19 (27.75)	2	2933.32	2
Świętokrzyski	633.32 (15.67)	15	2467.07	12
Warmińsko-mazurski	750.00 (17.42)	13	2398.00	15
Wielkopolski	1110.72 (23.02)	6	2610.82	7
Zachodniopomorski	987.72 (28.12)	8	2615.83	6
Average for Poland	1146.65 (9.99)	–	2866.04	–

Source: calculations based on the CSO data.

Table 2
The Moran's I and Geary's c for net wages in Poland

Moran's I					
Variables	I	E(I)	sd(I)	z	p-value*
Net wages	-0.120	-0.067	0.036	-1.494	0.068
Geary's c					
Variables	c	E(c)	sd(c)	z	p-value*
Net wages	1.102	1.000	0.092	1.113	0.133

*1-tail test.

Table 3
Estimation results

Variables	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Lnpotential	0.876*** (0.016)	0.874*** (0.013)	0.865*** (0.013)	0.626*** (0.006)	0.593*** (0.019)
Gender		-0.390*** (0.010)	-0.344*** (0.010)	-0.350*** (0.009)	-0.296*** (0.010)
Age		0.074*** (0.003)	0.068*** (0.005)	0.068*** (0.005)	0.061*** (0.004)
Age2		0.001*** (0.000)	0.001** (0.000)	-0.001*** (0.000)	0.001*** (0.000)
Education		-0.144*** (0.003)	-0.125*** (0.003)	-0.107*** (0.003)	-0.103*** (0.003)
Experience			0.008*** (0.003)	0.010*** (0.002)	0.008*** (0.002)
Experience2			0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Supervising			-0.330*** (0.014)	-0.291*** (0.013)	-0.273*** (0.012)
Agglomeration economies				-0.040*** (0.001)	-0.038*** (0.001)
Firm ownership					0.051*** (0.014)
maritalstatus_married		0.111*** (0.013)	0.090*** (0.014)	0.091*** (0.013)	0.080*** (0.013)
maritalstatus_widow(er)		0.094*** (0.036)	0.059* (0.035)	0.049 (0.032)	0.050 (0.031)
maritalstatus_divorced		0.083*** (0.028)	0.072*** (0.027)	0.022 (0.025)	0.024 (0.024)
region_dolnośląski				-0.131** (0.021)	-0.167*** (0.024)
region_kujawsko-pomorski				-0.253*** (0.025)	-0.272*** (0.028)
region_lubelski				-0.164*** (0.022)	-0.180*** (0.028)
region_lubuski				-0.293*** (0.025)	-0.337*** (0.033)
region_łódzki				-0.212*** (0.021)	-0.240*** (0.024)
region_małopolski				-0.077*** (0.026)	-0.109*** (0.028)
region_opolski				-0.471*** (0.027)	0.507*** (0.032)
region_podkarpacki				-0.203*** (0.023)	-0.239*** (0.030)
region_podlaski				-0.219*** (0.026)	-0.257*** (0.034)
region_pomorski				-0.044* (0.023)	-0.073*** (0.028)
region_śląski				-0.151*** (0.019)	-0.180*** (0.020)
region_świętokrzyski				-0.480*** (0.024)	-0.504*** (0.029)
region_warmińsko-mazurski				-0.224*** (0.023)	-0.251*** (0.031)

region_wielkopolski				-0.236***	-0.243***
				(0.020)	(0.022)
region_zachodniopomorski				-0.050***	-0.088***
				(0.025)	(0.031)
Firmsize_2					0.064
					(0.068)
Firmsize_3					0.104
					(0.068)
Firmsize_4					0.095
					(0.070)
Firmsize_5					0.132*
					(0.070)
Firmsize_6					0.148**
					(0.071)
Firmsize_7					0.149*
					(0.077)
Firmsize_8					0.137*
					(0.074)
Firmsize_9					0.162**
					(0.081)
Firmsize_10					0.206***
					(0.073)
firmsize_11-19					0.203***
					(0.064)
firmsize_20-49					0.252***
					(0.064)
firmsize_50-100					0.288***
					(0.064)
firmsize_101and more					0.357***
					(0.064)
firmsize_smaller that 11					0.074
					(0.072)
firmsize_more than 10					0.186***
					(0.067)
Sector dummies	NO	NO	NO	NO	YES
Constant	-3.521***	-3.855***	-3.095***	-0.538***	-1.355***
	(0.184)	(0.168)	(0.174)	(0.273)	(0.497)
Observations	9,825	9,825	9,825	9,825	9,825
R-squared	0.229	0.456	0.493	0.572	0.605

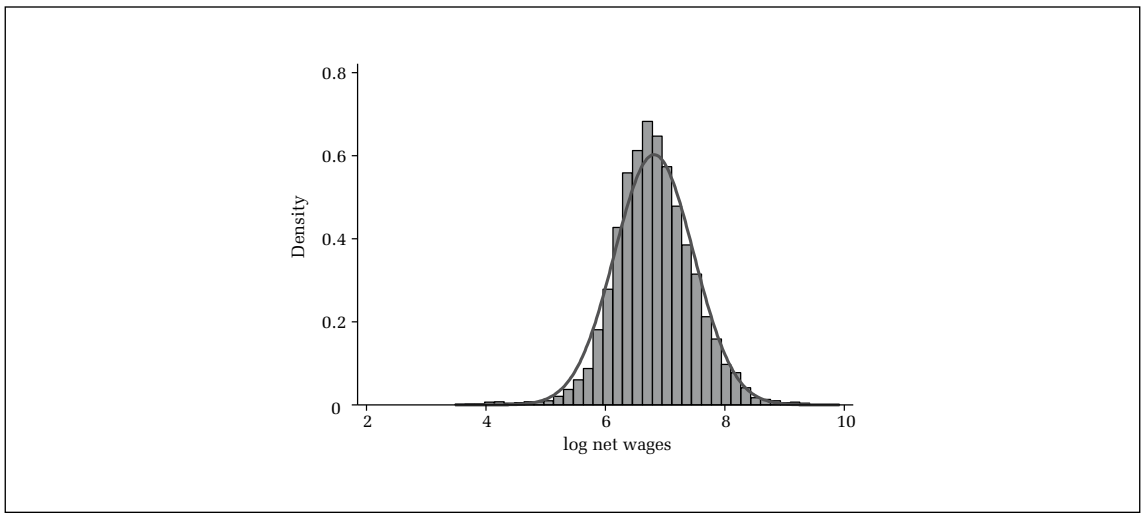
Notes:
– standard deviations in parentheses,
– ***, **, * denote statistical significance at 1%, 5%, and 10% confidence level, respectively.

Table 4
First stage estimation results

Variables in logs	Parameters
Employment	0.449*** (0.014)
Income	0.054*** (0.016)
Housing stock per capita	1.350*** (0.027)
Constant	1.003*** (0.081)
Observations	9,825
R-squared	0.930

Notes:
– standard deviations in parentheses,
– *** denotes statistical significance at 1% confidence level.

Figure 1
Histogram of log wages



Map 1
Spatial distribution of net wages in Poland