DIW Berlin
German Institute for Economic Research

Discussion Papers

545

Michal Myck
Leszek Morawski
Jerzy Mycielski

Employment Fluctuations and Dynamics of the Aggregate Average Wage in Poland 1996-2003

Berlin, January 2006

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Königin-Luise-Str. 5
14195 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

DIW Berlin

Discussion Papers 545

Michal Myck*

Leszek Morawski**

Jerzy Mycielski**

Employment Fluctuations and Dynamics of the Aggregate Average Wage in Poland 1996-2003

Berlin, January 2006

* DIW Berlin, Department of Public Economics, mmyck@diw.de

** Department of Economics, Warsaw University

Abstract:

The aggregate average wage is often used as an indicator of economic performance and welfare, and as such often serves as a benchmark for changes in the generosity of public transfers and for wage negotiations. Yet if economies experience a high degree of (nonrandom) fluctuation in employment the composition of the employed population will have a considerable effect on the computed average. In this paper we demonstrate the extent of this problem using data for Poland for the period 1996-2003. During these years unemployment in Poland almost doubled. We show that about a quarter of the growth in the average wage during this period could be contributed purely to changes in employment.

Key words: aggregate wage, employment dynamics, Poland

JEL classification: E24, J21, J31

Acknowledgments: The paper presents some of the results from the research project on "Aggregate wage dynamics in Poland". Financial support from the Polish Scientific Research Committee (Komitet Badań Naukowych) under grant number 2H02C06924 is gratefully acknowledged. We would like to thank Richard Blundell, Amanda Gossling and Howard Reed from the IFS for their advice in the early stages of the project's development. The usual disclaimer applies.

1. Introduction

The aggregate average wage is often understood and referred to as an indication of the performance of the economy and its dynamics as a reflection of changes in society's welfare. As such it is given a lot of public attention. As a result it is also used as a reference value for determining values of various social policy instruments.[1] It is taken as given that an increase in the average wage is a sign of positive developments in the economy, while it's stagnation reflects a general slowdown in economic development.

It has been recognised for a long time in economic literature that the reported dynamics of aggregate wages may not necessarily play the role it is commonly assigned, and the problem of meaningful aggregation of wages may be more complicated than it is usually perceived (Bills, 1985, Solon et al., 1994, Gossling et al., 1998, Meghir and Whitehouse, 1996).[2] The main issues complicating the interpretation of the aggregate wage as a simple indicator of welfare are:

- the fact that selection into and out of the sample of employees is not random,
- the structure of the employed population changes over time,
- the structure of the wage distribution may change over time.

A recent paper (Blundell et al., 2003) demonstrated that when corrected for these factors aggregate wage dynamics behave significantly differently from the simple average wage calculated for the employed population. While the measured aggregate wage in the UK over the early 1990's rises, the individual wages appear to be essentially flat. Clearly the interpretation of changes in the aggregate wage will be most difficult when the three forces complicating this interpretation undergo important changes. This will therefore apply especially strongly to countries with significant fluctuations in the rate of employment and more broadly to economies which undergo a rapid structural and institutional change.

[1] For example in Poland the level of income up to which national insurance contributions are paid is 30 times the average gross monthly wage from the previous year. Moreover, the computation of retirement and disability pension entitlements for those who become pensioners relates their earnings and contributions to average monthly gross wages. National insurance contributions paid by the self-employed also depend on average monthly gross wages. See Zdanowicz (2003).

[2] For an excellent survey on aggregation issues in economics see Blundell and Stoker (2005).

This paper aims to be a straightforward illustration of the factors driving the aggregate wage dynamics with an application to Poland. Our main goal is to demonstrate the importance of the three factors above and examine the extent to which they could have disguised what the average aggregate wage is supposed to reflect. We hope that our results will make the analysts who examine the trends in the average aggregate wage more careful in their interpretation of the reported series.

We use Polish micro-level data from two surveys: the Labour Force Survey (BAEL) and the Autumn Earnings Survey (AES). BAEL is a representative individual level survey (a rolling panel) collected quarterly with a principle focus on labour market status. Each quarter the survey collects information on about 50,000 individuals aged 15 and over. The AES is an annual survey (collected usually in September) which collects data on approximately 700,000 individuals at the company level and focuses on earnings information. The reason for using this joint set up is because wage information in BAEL is generally unsatisfactory.[3] At the same time the AES collects information only on employees and can't be used to analyse changes in participation patterns. Therefore we use the earnings information from the AES data and take advantage of the detailed labour market information from BAEL to study labour market dynamics.

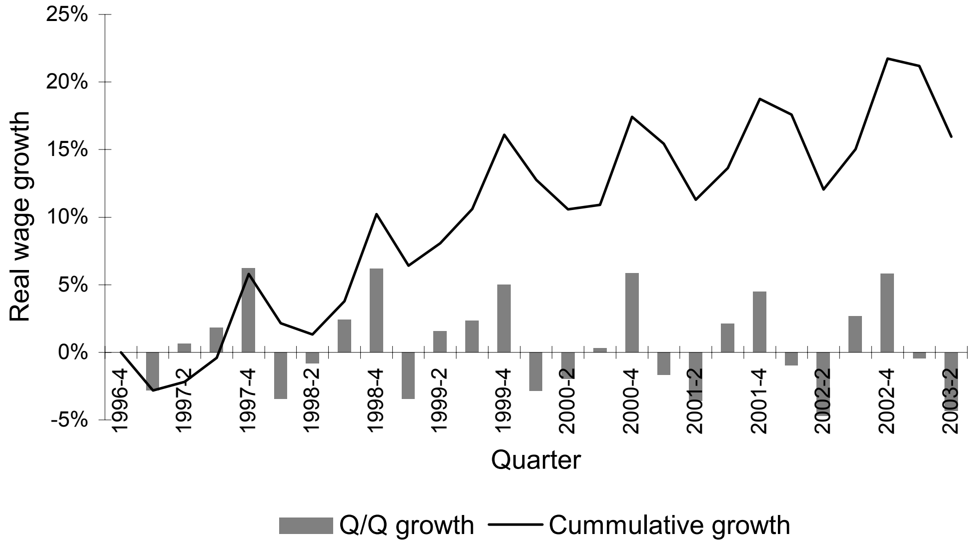
The detailed information on individual wages in the AES is used to generate a wage distribution for the BAEL sample. This distribution is then employed to demonstrate changes in the 'average wage' given the observed trends on the Polish labour market in the period from the last quarter of 1996 to the middle of 2003. This time was a period of significant changes in employment levels in Poland. Registered unemployment in this period initially fell from 13.2% in September 1996 to 9.5% in August 1997, but then rose to reach 18.3% in April 2003. Below we look at the effects this employment variation could have had on the interpretation of the aggregate average wage.[4] Because of institutional changes it is impossible to present a

[3] As we show in Mycielski et al. (2005) the wage information in BAEL is substantially incomplete. Moreover the AES collects information on gross and not on net wages (as is the case in BAEL). This makes it more comparable with the official average wage statistics.

[4] The published information on the aggregate average wage in Poland is based on monthly surveys conducted by the Central Statistical Office. They cover all financial and non-financial enterprises employing at least 50 people and a representative sample of 10% of firms with from 10 up to 50 employees. Wages in very small firms are estimated on the basis of trends from previous years.

consistent series of the aggregate average wage for Poland, as the definition of the "gross wage" changed in January 1999 with the introduction of the pension reform. However, it is possible to calculate separate rates of growth of the aggregate average wage for the period before January 1999 and for the period after. By assuming the rate of growth between Q4 of 1998 and Q1 of 1999 to be the same as that between Q4 1997 and Q1 1998, we can construct a cumulative wage growth over the period we examine. This is shown on figure 1. We can see that the cumulative real wage growth from the beginning to the end of the period is about 16%, but it reaches its peak in the fourth quarter of 2002 at 21.7%.

Figure 1. Growth of real aggregate average wage in Poland: 1996-Q4 – 2003-Q2



Source: authors' computation based on GUS statistics.
Notes: Growth rate of real earnings between Q4-1998 and Q1-1999 assumed the same as a year earlier. Growth rates computed on the basis of CPI-indexed nominal wages.
Q/Q – change from previous quarter.

We start by outlining our methodology in Section 2. In Section 3 we describe our data and present sample selection criteria together with basic descriptive statistics for the data used in the study. Section 4 presents the constructed 'base' wage distribution, while in section 5 we show results of aggregate wages dynamics for the period from the fourth quarter of 1996 to the second quarter 2003 under the assumption of constant base wage distribution. Conclusions follow in section 6.

Average wages are calculated as an arithmetic mean with total wage fund divided by the number of employed people which is in some cases adjusted for number of working hours.

2. Methodology

Our aim is to illustrate the most important determinants of changes in the 'average wage' in a straightforward but realistic fashion. This, together with the limitations imposed by the availability of data, implies that the analysis will need to be based on several simplifying assumptions, but it should nevertheless serve as a useful illustration of the problem and an interesting application to Polish data.

The average wage is usually calculated simply as:

$$\varpi_t = \frac{\sum_{i=1}^{I_t} w_{it}}{I_t} \tag{1}$$

where: where w_{it} is the observed wage of individual (*i*) at time (*t*) and I_t is the sample of people employed at time (*t*) (i.e. people for whom we observe a wage). Notice that this definition disregards the part time and full time work and treats part and full time employees equally.

The formula obviously relates only to the wages of those who are employed and is calculated as a simple arithmetic average. This means that although it may be informative of the wage level at a particular point in time, its changes may be difficult to interpret, especially if the population of employees (I_t) changes between (*t*) and (*t*+*s*). Because changes in the population of employees, driven by the economic cycle, demand for labour and by individual labour supply decisions, are not random, at different points in time people from different sections of the wage distribution will leave the sample or join it. The analysis of changes in the average wage should therefore take into account the changing composition of the employee population. At different points in the economic cycle people may be fired or hired and/or decide to leave employment or take up a job. This of course presents a difficulty because wages for the non-employed population are not observed. The average wage continues to be calculated in the same way, although the sample I_t changes. One therefore calculates the average wage for different people at each *t*.

How different would be the situation if we could observe wages for the entire population? If everyone had a wage 'label' which would respond to market demand factors, we would then be able to calculate a meaningful measure of the aggregate wage for the entire population at each (*t*). This would not help with the two other problems we mentioned in the introduction, but the issue of non-random selection would be taken into account in the calculation of the aggregate wage. Yet, clearly the assumption that we can observe wages of all working age individuals is wrong and in reality we can only observe wages of those in work. So the problem we have to address is how we can use this information to learn about the effect of the three factors complicating the interpretation of the aggregate wage dynamics?

Let us go back to the assumption that we can observe individual wages for the entire working population, and let's also assume that the working population is made only of the 'employed' and the 'non-employed'. The average wage could be calculated for this population as in formula (1) taking account of the wages of the employees. We could also proceed in a simplified manner, by dividing the distribution of employees' wages into a number of intervals (c) and then calculating the average as:

$$\varpi'_t = \frac{\sum_{c=1}^{C} (\overline{w}_{ct} * \eta_{ct})}{\sum_{c=1}^{C} \eta_{ct}} \tag{2}$$

where *n* is the number of employees in each wage interval *c* and $\overline{w}_{ct}$ is the arithmetic average wage for that interval *c* of the employed population. In the limit (when each interval represents a single individual) this formula is identical to (1), but this different representation of the way the average can be calculated becomes useful when we want to account for the entire working population.

Assuming that the values of wages are conditional only on the observed characteristics of individuals and are independent from the current employment state the average wage formula (2) can be written in terms of the probabilities of being observed as employed:

$$\varpi''_t = \frac{\sum_{c=1}^{C}\left(\tilde{w}_{ct} * \sum_{i=1}^{Nc} p_{ict})\right)}{\sum_{c=1}^{C}\left(\sum_{i=1}^{Nc} p_{ict}\right)} \tag{3}$$

where *Nc* is the number of employed and non-employed people in a wage interval c, and p_{ict} is the individual probability of being employed. $\tilde{w}_{ct}$ is the average wage within the interval calculated for the employed and the non-employed individuals. It is clear that if we assign the observed probabilities of being employed or not being employed (i.e 1 to those observed as employed and 0 to those observed as non-employed) then formula (3) is identical to (2) (provided that the intervals c are small enough so that the distribution of wages within the intervals can be treated as uniform). Representation (3) is however more flexible from the point of view of our illustrative aim as it allows to examine what happens to the aggregate wage not only if wages of the employees change but also if the probability of being employed p_i changes. We can also present formula (3) at the individual level (for the case where p_i is individual probability of being employed, i.e. there is only one individual per interval):

$$\varpi'''_t = \frac{\sum_{i=1}^{Nc}(\tilde{w}_{it} * \hat{p}_{it}))}{\sum_{i=1}^{Nc}\hat{p}_{it}} \tag{4}$$

Notice that we can replace p_i with $\hat{p}_i$ - the expected probability of being employed and in this way analyse how changes in $\hat{p}_i$ over time would have influenced the aggregate average wage controlling for the entire (expected) wage distribution. The wage measure included in formula 4 is $\tilde{w}_{it}$ - the assumed known individual wage for individual (*i*). In our exercise this will be replaced with $\hat{w}_{it}$ - the *expected* wage measure of individual (*i*). Formula 4 Below we present an analysis for Poland for the period from fourth quarter of 1996 to the second quarter of 2003 using this approach. We demonstrate the extent to which changes in employment probability of people with certain characteristics could have affected the dynamics of the average wage.

3. Data

As we pointed out in the introduction the analysis is based on the combination of the AES and BAEL data. We use AES collected in September 1996 and on the basis of this dataset generate the expected (gross) wage distribution for the BAEL sample collected in Autumn 1996 (referred to as the 'BAEL base sample' below). Following this exercise we estimate employment probabilities in BAEL datasets over the years 1996-Q4 to 2003-Q2 and use these to generate expected employment probabilities in 'future' BAEL years, which are calculated for the BAEL base sample.[5]

Where possible the same sample selection criteria are applied to the AES and BAEL datasets. The most important selection criteria are:

- we drop people aged less than 18 and over 60 (in both samples),
- we drop the self employed, those who help in family business and full time students (in BAEL only).

In table 1 we present the basic descriptive statistics for the AES sample and the BAEL base sample (after applying selection criteria). The table also includes descriptive statistics for a sub-sample of the BAEL base sample including only those who are employed in firms employing more than 5 employees. This is the closest we can get to mimic the criteria applied to the creation of the AES sample. The Autumn Earnings Survey collects information only on employees employed in firms with more than 9 employees. We can see that as far as the proportion of higher educated and those with secondary vocational training the selected BAEL sample and the AES sample is very similar. However the proportion of men among the BAEL employees is higher than among the AES employees, and the proportion of those with primary education is lower. These differences will most likely plead to differences in the earnings distributions generated for AES and for BAEL, but they are not very important as far as the analysis of changes in aggregate wages is concerned.

[5] The BAEL data has undergone a significant "transformation" in the period covered in this paper. For detailed documentation on changes in the data see Morawski, Mycielski, Myck (2005).

4. Computing expected wages and employment probabilities

Table 2 presents the results of the wage equation estimation run as truncated regression (on log monthly wage) on AES.[6] Log wages (of 663,621 individuals) are regressed on an age polynomial, education level dummies, region (49 pre-1999 voivodship), a male dummy and interactions of: the male dummy with regional dummies and with education dummies. The choice of these variables is constrained on the one hand by the availability of more information on individuals in the AES dataset, and on the other hand on the need to have the same variables in both the AES and the BAEL base sample.[7] For presentation reasons table 2 does not include the coefficients on the regional dummy variables and the interactions.[8]

Table 1. AES and BAEL sample - descriptive statistics

	AES – 1996[1]	BAEL – 1996-Q4[2]	BAEL – 1996-Q4[3] (AES selection)
Sample size	676,360	30,048	16,667
Proportion of men	50.9%	47.6%	54.5%
Education:			
- higher	15.5%	10.0%	14.6%
- secondary academic	6.9%	6.9%	6.5%
- secondary vocational	30.5%	27.7%	31.7%
- vocational	30.0%	35.8%	35.2%
- primary or none	17.0%	19.6%	11.9%
Age – men	38.7	38.5	37.3
Age - women	39.3	38.9	38.5

Source: authors' calculation on the basis of AES-1996 and BAEL-1996-Q4.
Notes: 1 - individuals employed in companies with more than 9 employees.
2 - employed and non-employed individuals.
3 - individuals employed in companies with more than 5 employees.

The results are not very surprising. Wages are higher for older and better educated people. The coefficient on the male dummy variables is positive, but men get lower returns to age and education.

[6] The truncation is made at the level of the National Minimum Wage (325PLN) on the left hand side and at the 99th centile of the distribution (2706.80PLN) on the write hand side of the distribution. See figure A1 in the Appendix for the shape of the lower end of the distribution before the truncation.
[7] For example we could not use the work experience information from AES as such information is not available in BAEL.
[8] A significant majority of those coefficients is statistically significant. the full set of results is available from the authors on request.

This estimation is used to generate a distribution of expected wages in AES and in the BAEL base sample (shown on figure 2a and 2b).[9] The average expected wage for men and women in AES is 788.00 and 651.30 respectively, and BAEL base sample these numbers are: 780.40 and 611.40. Figure 2b – plotted for the BAEL base sample includes both the employed (in big and in small firms) as well as the non-employed people. Thus the difference between the two distributions should not be surprising. It is noteworthy that the most pronounced difference between the two charts is at the lower end of the wage distribution which is the effect of the fact that people with lowest human capital (expressed in the expected wage) are more likely not be employed and thus not be observed in the AES sample. The same argument would explain the difference in the average level of expected wages for both men and women.

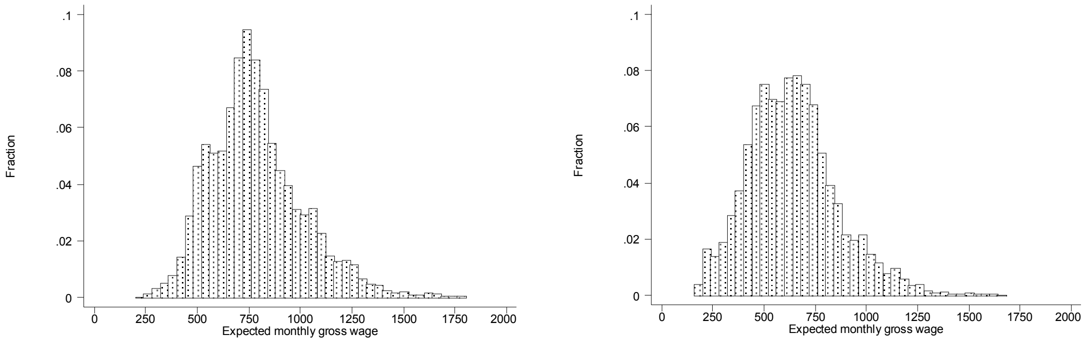
Table 2. Summary results of truncated wage regression in AES, 1996

Dependent variable: log monthly gross wage	**Coeff.**	**St. error**	**Significance level**
Age	0.4815	(0.0206)	***
Age^2	-0.0164	(0.0008)	***
Age^3	0.0003	(0.0000)	***
Age^4	0.0000	(0.0000)	***
Education (base cat.: primary)			
- higher	0.6741	(0.0031)	***
- secondary academic	0.4015	(0.0034)	***
- secondary vocational	0.3881	(0.0028)	***
- vocational	0.0647	(0.0033)	***
Male dummy	0.6253	(0.2394)	***
Age*male dummy	-0.0397	(0.0262)	
Age^2*male dummy	0.0022	(0.0010)	**
Age^3*male dummy	0.0000	(0.0000)	***
Age^4*male dummy	0.0000	(0.0000)	
Education * male dummy			
- higher * male dummy	-0.1871	(0.0041)	***
- secondary academic * male dummy	-0.1995	(0.0059)	***
- secondary vocational * male dummy	-0.1576	(0.0037)	***
- vocational * male dummy	0.0134	(0.0040)	***
Regional dummies		included	
Regional dummies*male dummy		included	
Sigma	0.4085	(0.0005)	***
Number of uncensored observations:		653,250	
Number of censored observations:		10,371	
Log likelihood		-240134.61	

Source: authors' calculations on the basis of AES 1996.
Notes: Observations truncated at the National Minimum Wage (325zl) and at the top centile of the wage distribution. *** - significant at 1%, ** - significant at 5%.

[9] Expected wages are computed as: $\hat{w}_i = \exp(l\hat{w}_i) * \exp(1/2\hat{\sigma}^2)$. See Blundell et al. (2003).

Figure 2. Expected gross monthly wage distributions – AES and BAEL base sample



2a - Expected wage distribution in AES

2b - Expected wage distribution in the BAEL base sample

Source: authors' calculations on the basis of BAEL 1996-Q4 and AES 1996.

The next step in the methodology is the estimation of employment probability models for BAEL samples over the period 1996-Q4 to 2003-Q2. Table A1 in the appendix presents the full set of results from the employment probit models run on 25 quarterly BAEL samples.[10] In Figure 3 we present the dynamics of the mean, median and 90th and 10th percentile expected employment probabilities calculated for the BAEL base sample on the basis of employment probit equations run on BAEL samples from 1996-Q4 to 2003-Q2.[11] We thus calculate employment probabilities for people in the BAEL base sample *as if* the conditions they were subjected to were imported from future years. Employment probability changes significantly during the period from 1996-Q4 to 2003-Q2 and the expected probability figures reflect the overall trends of the economy. From the point of view of the analysis presented in this paper it is important that the fall in the probability of being employed is not uniform. We can see that the 90th percentile probability falls only by about 6 percentage points, while the median and the 10th percentile probability fall by about 12 percentage points. Thus people at the lower end of the probability spectrum are much less likely to be employed at the end of the period than at the beginning of the period then people at the higher end of the spectrum.

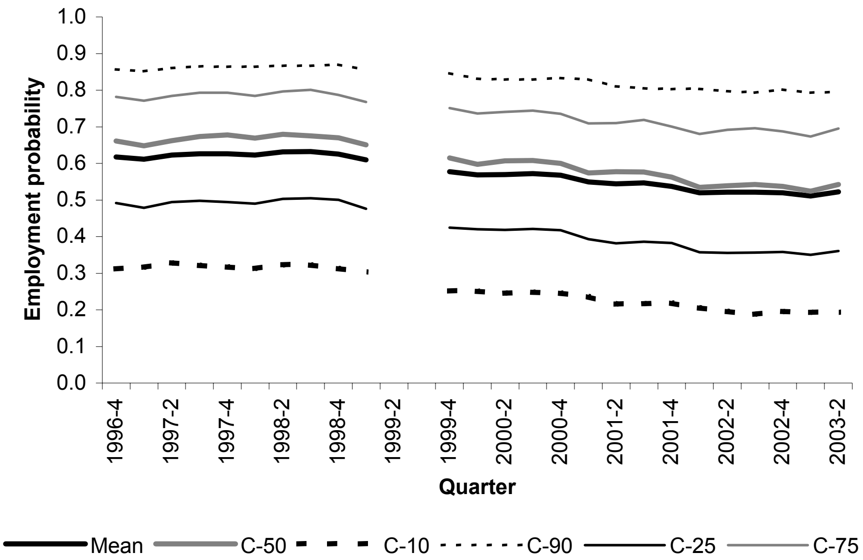
[10] The sample selection criteria in these cases were identical to those applied to the BAEL base sample.
[11] Note that BAEL was discontinued in 1999 and there is no data available for the second and third quarter of 1999.

5. Aggregate average wage dynamics 1996 – 2003

Since we now have a wage measure for every individual in the BAEL base sample, and measures of his/her probability of being employed (at different points in time), we can employ formulas from section 2 to derive the aggregate average wage dynamics for our sample.

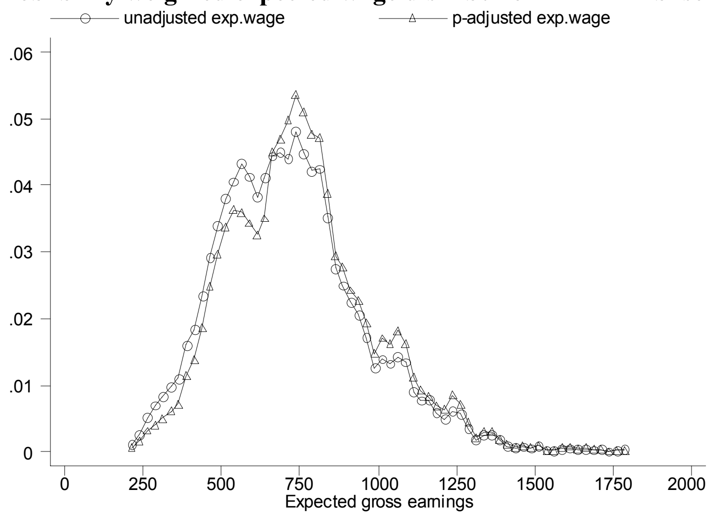
We start by presenting the difference probability weighting makes to the distribution of earnings in the BAEL base sample. The wage distribution is divided in 60 intervals of 25PLN. They are constructed in such a way that each of the expected wage measure falls into one of these intervals. If the frequencies with which people fall into these intervals are unweighted, then a histogram of these frequencies will (very closely) reflect the wage distribution from figure 2b. By applying employment probability weights to these frequencies we in a way "fire" people from these intervals. Since these probabilities are higher for those at the lower end of the earnings distribution, the distribution will shift thus affecting the calculated aggregate average wage. The unadjusted and the adjusted expected wage distributions are shown on figure 4. It is interesting to note that the probability-adjusted distribution is looks more alike the expected wage distribution from AES shown in figure 2. The aggregate average unadjusted wage is 719.92zl and the average rises to 754.29 once we adjust for employment probabilities. Probability weighing thus increases the computed average by 4.8%.

Figure 3. Trends in expected employment probability. BAEL – base sample



Source: authors' calculations on the basis of BAEL 1996-Q4 – 2003-Q2;
Notes: probabilities calculated on the basis of estimates on data from a specific quarter which were applied to BAEL base sample.

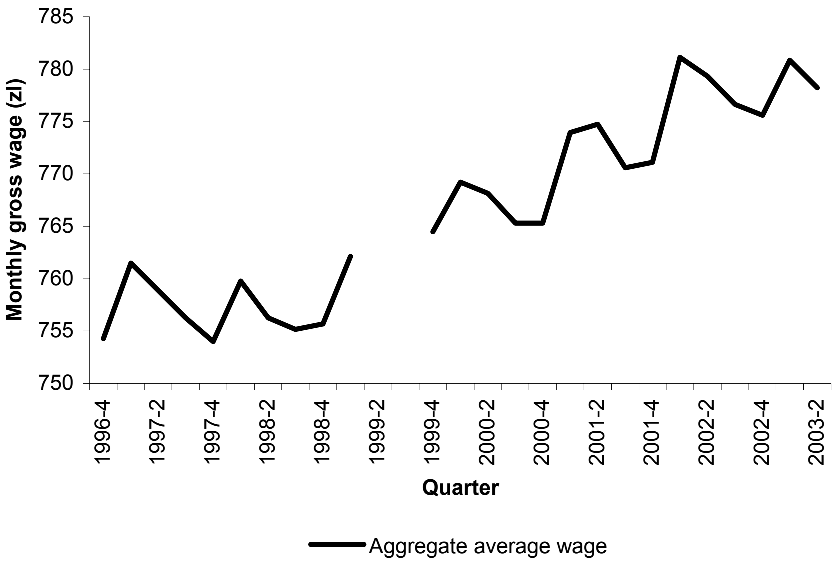
Figure 4. Probability weighted expected wage distribution in BAEL base sample



Source: authors' calculations on the basis of BAEL 1996-Q4 and AES 1996.

In the same way we employ formulas (3) and (4) to compute aggregate average wages for our sample using employment probabilities computed on the basis of BAEL 1997-Q1 to 2003-Q2. By changing the probability weights we are able to examine how changes in the probability of employment affect the computed aggregate average wage at the time when the *underlying* wage distribution (i.e. the distribution of expected individual wages) remains unchanged. The only thing that affects the computed average wage is the probability that someone is in the sample of employed people or not. The result of this exercise is plotted on figure 5.

Figure 5. Probability of employment and dynamics of aggregate average wage



Source: authors' calculations on the basis of BAEL 1996-Q4 and AES 1996.

Changes in employment probability seem to have a very strong (and seasonal) effect on the aggregate average wage measure. In a scenario where individual wages remain unchanged and we only change the probability of being employed the computed aggregate average wage rises from 754.27zl in the fourth quarter of 1996 to 778.23zl in the second quarter of 2003, a change of 3.2%. The average wage is lowest in the fourth quarter 1997 (754.00) and highest in the first quarter 2002 (781.11zl) – here the average wage rises by 3.6% in the space of three years just because of changes in the probability of employment. The highest difference in the computed average from quarter to quarter is between the fourth quarter 2001 and the first quarter of 2002. Here the difference is 1.3% - driven entirely by changes in individual employment probabilities.

6. Conclusion

We have presented an exercise of simulating changes in the aggregate average wage which result purely from changes in the structure of employment with the underlying distribution of individual wages remaining unaffected. This is a simple, but to our knowledge so far not implemented, way of decomposing changes in the aggregate average wage into those which result from actual changes in productivity and those which are sole reflections of changes in the composition of the employed population.

The methodology was applied to Polish data on earnings and employment from the Autumn Earnings Survey and the Polish LFS (BAEL) respectively.

The analysis shows that changes in employment in Poland must have had a significant effect on the observed dynamics of aggregate average wage in the second half of 1990s and the first few years of the twenty first century. Our estimation suggests that this effect was in the range of 3.5% and points to the important fact that seasonal changes in employment may significantly affect the average wage between two consecutive seasons.

Our results are most likely to be the lower bound estimates of the effect of employment on average wage dynamics. This is because throughout the analysis we have assumed that wages are only determined by observed characteristics. This is clearly a very strong assumption. If it does not hold, and unobserved heterogeneity affects both wages and employment (and there is positive correlation between these effects) then the actual effect of changes in employment on the dynamics of aggregate average wage would be even stronger.

Aggregate real wage increased in Poland by about 16% in the period covered by our analysis. This would mean that almost a quarter of the real growth in the average wage could be attributed purely to changes in the structure of employment between the end of 1996 and mid 2003. This degree of overestimation of actual (productivity related) changes in wages in the economy by using the “average”, must have had significant effects on the economy. The published wage statistics are often used as a benchmark for wage negotiations. Moreover, the overestimation of changes in real wages must have head significant fiscal consequences, given the use of the “average wage” as a reference point for several fiscal policy instruments.

References

Bils, M. J. (1985). 'Real Wages Over the Business Cycle: Evidence from Panel Data', *Journal of Political Economy*, 93, pp. 666-689.

Blundell, R., Reed, H. and Stoker, T. (2000), 'Interpreting Aggregate Wage Growth', *IFS Working Paper W99/13* (revised version), London: IFS, www.ifs.org.uk.

Blundell, R., Reed, H. and Stoker, T. (2003), 'Interpreting Aggregate Wage Growth: The Role of Labor Market Participation', *American Economic Review*, 94(3), pp.1114-1131.

Blundell, R. and T. Stroker (2005), 'Heterogeneity and Aggregation', *Journal of Economic Literature*, 28, pp. 347-391.

Gosling, A., Machin, S. and Meghir, C. (2000), 'The Changing Distribution of Male Wages in the UK', *Review of Economic Studies*, 67(4), pp. 635-66.

Heckman, J. and Sedlacek, G. (1990), 'Self-Selection and the Distribution of Hourly Wages', *Journal of Labor Economics,* 8, S329-S363.

Meghir, C. and Whitehouse, E. (1996), 'The Evolution of Male Wages in the UK: Evidence from Micro-Data', *Journal of Labour Economics*, 14(1), 1-25.

Morawski,L., Mycielski, J. and Myck, M. (2005), The Polish Labour Force Survey – BAEL 1994 –2004 – database documentation, Warsaw University – Department of Economics, *mimeo*.

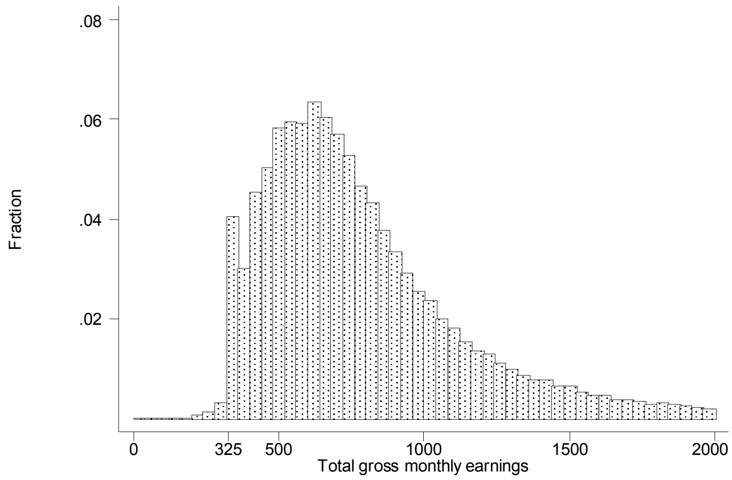
Mycielski, J., Morawski, L. and Myck, M. (2005), Wage data in the Polish LFS – the increasing problem of missing data. Warsaw University – Department of Economics, *mimeo*.

Solon, G., Barsky, R. and Parker, J.A. (1994), 'Measuring the Cyclicality of Real Wages: How important is Composition Bias', *Quartery Journal of Economics*, 109, pp.1-25.

Zdanowicz, J. (2003), Wskazniki przecietnego wynagrodzenia liczone przez Glówny Urzad Statystyczny, Warsaw University – Department of Economics, *mimeo*.

Appendix

Figure A1. Lower end of the wage distribution in AES 1996 – before truncation



Notes: 325PLN was the value of the National Minimum Wage in Poland in September 1996.

Table A1. Results from employment probability models - winter quarters 1996-2002 (and summer quarter 2003)

	Q4 - 1996		Q4 - 1997		Q4 - 1998		Q4 - 1999		Q4 - 2000		Q4 - 2001		Q4 - 2002		Q2 - 2003	
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
Age	-0.4272	(0.1667)	-0.4960	(0.1646)	-0.2321	(0.1650)	-0.4085	(0.1773)	-0.3350	(0.1810)	0.2383	(0.2044)	-0.2235	(0.2038)	0.7634	(0.2179)
Age^2	0.0146	(0.0069)	0.0168	(0.0068)	0.0066	(0.0068)	0.0170	(0.0073)	0.0130	(0.0075)	-0.0105	(0.0083)	0.0097	(0.0083)	-0.0301	(0.0088)
Age^3	-0.0002	(0.0001)	-0.0002	(0.0001)	0.0000	(0.0001)	-0.0003	(0.0001)	-0.0002	(0.0001)	0.0002	(0.0001)	-0.0001	(0.0001)	0.0006	(0.0002)
Age^4	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)
- higher	1.3735	(0.0472)	1.4861	(0.0475)	1.5206	(0.0461)	1.5371	(0.0499)	1.5433	(0.0486)	1.5037	(0.0497)	1.5222	(0.0480)	1.5406	(0.0488)
- post-secondary	0.9788	(0.0544)	1.1178	(0.0558)	1.1809	(0.0567)	1.1748	(0.0608)	1.0626	(0.0586)	1.0051	(0.0602)	1.0176	(0.0598)	1.0474	(0.0602)
- secondary vocational	0.7274	(0.0330)	0.7802	(0.0334)	0.8450	(0.0338)	0.8199	(0.0378)	0.7954	(0.0376)	0.7907	(0.0392)	0.7799	(0.0396)	0.8080	(0.0403)
- secondary academic	0.5526	(0.0409)	0.6499	(0.0410)	0.6245	(0.0415)	0.7162	(0.0457)	0.6633	(0.0451)	0.6266	(0.0483)	0.7151	(0.0481)	0.6821	(0.0483)
- vocational	0.2985	(0.0327)	0.3216	(0.0330)	0.3930	(0.0333)	0.3203	(0.0379)	0.3453	(0.0375)	0.3524	(0.0386)	0.3350	(0.0390)	0.3017	(0.0398)
Male	-16.628	(2.0398)	-17.697	(2.0663)	-16.136	(2.0889)	-14.057	(2.2412)	-15.312	(2.2435)	-10.645	(2.5560)	-15.187	(2.5706)	-3.2370	(2.6926)
Town 10k	0.1001	(0.0198)	0.0604	(0.0198)	0.0482	(0.0200)	0.0758	(0.0217)	0.0356	(0.0213)	0.0506	(0.0217)	0.0100	(0.0216)	0.0538	(0.0211)
Town 100k	0.1427	(0.0206)	0.1371	(0.0207)	0.1494	(0.0208)	0.1769	(0.0220)	0.1947	(0.0216)	0.2054	(0.0223)	0.1131	(0.0221)	0.1669	(0.0222)
Age * male	1.8360	(0.2356)	1.9519	(0.2382)	1.7483	(0.2400)	1.5219	(0.2573)	1.6259	(0.2570)	1.0538	(0.2883)	1.5727	(0.2889)	0.2032	(0.3011)
Age^2 * male	-0.0681	(0.0098)	-0.0723	(0.0099)	-0.0630	(0.0099)	-0.0551	(0.0106)	-0.0581	(0.0106)	-0.0337	(0.0117)	-0.0552	(0.0117)	0.0007	(0.0122)
Age^3 * male	0.0011	(0.0002)	0.0011	(0.0002)	0.0009	(0.0002)	0.0008	(0.0002)	0.0009	(0.0002)	0.0004	(0.0002)	0.0008	(0.0002)	-0.0002	(0.0002)
Age^4 * male	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)
- higher * male	-0.2587	(0.0713)	-0.1844	(0.0750)	-0.2143	(0.0742)	-0.2937	(0.0765)	-0.3033	(0.0746)	-0.0439	(0.0768)	-0.1462	(0.0728)	-0.1964	(0.0726)
- post-secondary * male	0.0168	(0.1373)	-0.4795	(0.1228)	-0.5285	(0.1231)	-0.3609	(0.1368)	-0.2462	(0.1295)	-0.1550	(0.1230)	-0.0794	(0.1198)	-0.3362	(0.1148)
- secondary vocational * male	-0.1132	(0.0492)	-0.1292	(0.0497)	-0.1274	(0.0505)	-0.0219	(0.0548)	-0.0613	(0.0543)	-0.0068	(0.0566)	0.0388	(0.0563)	-0.0130	(0.0573)
- secondary academic * male	-0.1315	(0.0824)	-0.3770	(0.0808)	-0.2870	(0.0813)	-0.0943	(0.0843)	-0.1014	(0.0813)	0.0727	(0.0878)	0.0249	(0.0868)	-0.0511	(0.0820)
- vocational * male	0.1105	(0.0450)	0.1322	(0.0458)	0.0330	(0.0463)	0.1789	(0.0514)	0.1172	(0.0511)	0.1683	(0.0528)	0.1448	(0.0529)	0.2006	(0.0540)
regional dummies	included		included		included		included		included		included		included		included	
Number of observations	30048		30094		29738		25222		26002		24316		24692		24840	
Log likelihhod	-16665.30		-16391.00		-16247.80		-14303.80		-14866.90		-14150.40		-14385.00		-14465.30	

Source: authors' calculations on the basis of BAEL data.