Driving forces behind the sectoral wage costs differentials in Europe*

Camille Logeay**
Macroeconomic Policy Institute
(IMK), Duesseldorf
Camille-Logeay@boeckler.de

Sabine Stephan
Macroeconomic Policy Institute
(IMK), Duesseldorf
Sabine-Stephan@boeckler.de

Rudolf Zwiener
Macroeconomic Policy Institute
(IMK), Duesseldorf
Rudolf-Zwiener@boeckler.de

This version: July 18, 2008

---

*We thank Simon Sturm for helpful comments, the usual disclaimer applies.
**Corresponding author, Camille Logeay, IMK, Hans-Boeckler-Str. 39, D-40 476 Duesseldorf, Germany; phone +49/211/7778 334, fax +49/211/4 7778 334.
Abstract

In 2004, Eurostat starts publishing new figures on hourly wage costs for all European countries. These figures are new in several respects: It is the first time that internationally comparable hourly figures on wage costs are available covering a quite important time period (1995-2005), so that not only cross-country comparisons but also dynamic analyses are possible. Furthermore, these figures are fairly detailed at the sectoral level, therefore allowing for inter-sectoral comparisons. Concerning Germany, the Eurostat statistics provide quite unexpected insights; the gap between wage costs in the manufacturing sector and the (private and business) services sector is much larger than in other countries. This study aims at giving some explanations. According to theory, various explanations are possible. First, the neo-classical theory emphasises factors affecting or indicating the level of individual productivity, as well as firm or sectoral productivity; indicators corresponding to this approach are tested. Second, dropping the assumption of perfect competition on both labour and goods markets allows for other factors (mark-up, market power) to influence the wage costs levels; these potential determinants are also tested. Finally, we think that the structure of demand (driven by domestic or foreign demand) could also have a major impact on wages in the industry and the services sector and indeed this factor seems to play an important role. This paper is structured as follows: First, the new Eurostat statistics is presented focussing on some interesting descriptive results. In the second section, we present a list of potential determinants of wage differentials between the industry and the services sector derived from theory and literature. A bivariate analysis (correlation) is then performed and conclusions are drawn. In a third step, a multivariate analysis (panel estimation) is performed. The final section concludes.

JEL: J31, C23, E24

Keywords: Wage differentials, Europe, sectoral level, macroeconomic panel.
1 Introduction

In 2004, Eurostat starts publishing new figures on hourly wage costs for all European countries. These figures are new in several respects: It is the first time that internationally comparable hourly figures on wage costs are available covering a quite important time period (1995-2005), so that not only cross-country comparisons but also dynamic analyses are possible. Furthermore, these figures are fairly detailed at the sectoral level, therefore allowing for inter-sectoral comparisons.

Concerning Germany, the Eurostat statistics provide quite unexpected insights; the gap between wage costs in the manufacturing sector and the (private and business) services sector is much larger than in other countries. This study aims at giving some explanations. According to theory, various explanations are possible. First, the neo-classical theory emphasises factors affecting or indicating the level of individual productivity, as well as firm or sectoral productivity; indicators corresponding to this approach are tested. Second, dropping the assumption of perfect competition on both labour and goods markets allows for other factors (mark-up, market power) to influence the wage costs levels; these potential determinants are also tested. Finally, we think that the structure of demand (driven by domestic or foreign demand) could also have a major impact on wages in the industry and the services sector and indeed this factor seems to play an important role.

This paper is structured as follows: First, the new Eurostat statistics is presented focussing on some interesting descriptive results. In the second section, we present a list of potential determinants of wage differentials between the industry and the services sector derived from theory and literature. A bivariate analysis (correlation) is then performed and conclusions are drawn. In a third step, a multivariate analysis (panel estimation) is performed. The final section concludes.
2 Overview of the literature

2.1 Theory: several competing explanations

There are different theories explaining persisting wage differentials. The neo-classical tradition focuses on factors affecting or indicating the level of individual productivity as well as firm and sectoral productivity respectively. Theories dropping the assumption of perfect competition on both labour and goods markets consider additional determinants (mark-up, market power) influencing wage cost levels. We think, however, that the structure of demand (domestic versus foreign demand) could also play an important role in determining wages in specific sectors.

Schramm (2004, chap. 3) distinguishes three different explanations for cross-sectoral wage differentials: First, the neo-classical approach assuming perfect competition on the labour market and explaining wage differentials due to different labour/job quality or compensation for non-monetary remunerations. Second, the efficiency wage theory broadens the neo-classical view assuming that wages above the neo-classical equilibrium give workers an incentives to work in a more efficient manner. In doing so they produce a rent which can be divided among employers and employees (rent-sharing). Moreover, employers act reasonable if they pay more than the equilibrium wage because of moral hazard problems. The third approach focuses on non-competitive features on the labour and goods markets due to institutional set-ups that give rise to rent-sharing for both sides, independently from the productivity level. Sociological aspects are also put forward to explain the persistence of the differences over the business cycles.

Schettkat (2006, chap. 2) quotes the same explanations and puts more emphasis on the assumption of a monopsonistic labour market (Manning 2003). In this model (contrary to the efficiency wage) the actual wage is fixed below the optimal level. A lot of empirical case studies, esp. in relation to the introduction of a minimum wage in the UK (Metcalf 2007), show that several sub-branches of the service sector are better modeled as a monopsony rather than within a perfect competition framework (Card and Krueger 1994, Dickens and Katz 1986, Card 1996, Machin and Manning 2002). Thus not only productivity differences but also the structure of the labour market could explain wage differentials across sectors.

The efficiency wage theory assumes that the wage level directly affects workers’ ef-
forts and that the information about the strength of this effort is asymmetric; workers know, employers can only guess (Shapiro and Stiglitz 1984, Krueger and Summers 1988). Consequently, moral hazard problems occur. This is the reason why employers have incentives to take workers’ interests into account. Wages above the optimal wage level (without efficiency wage consideration) produce involuntary unemployment but reduce turn-over (hiring/firing costs) and provide incentives for high-productive workers to apply and stay in the firm and to increase their work-effort. Therefore, cross-sectional wage differences can be due to different productivity/qualifications of workers as stated in the neo-classical theory; but they can be modified according to the social consensus – if society prefers a more equal income distribution, we would face a smaller wage dispersion (due to a higher wage compression) than the one occurring according to productivity/quality differences in the specific sectors.

Nowadays, institutional factors are also put forward; especially the presence of a coordinated wage bargaining system should reduce wage dispersion across sectors and qualifications. Thus, a high union density or a high coverage of union agreements should go hand in hand with a smaller wage differential (Layard, Nickell, and Jackman 1991, “the battle of markups”). The market power of trade unions, however, probably depends on the firm size; Very small firms could be able to resist trade unions’ demands more easily than bigger ones (where strikes could be more likely and costly). Not only the size of the firms may play a role but also the amount of the rent the owner receives. A higher rent could encourage workers to organize in a trade unions in order to force the owner to share it. In the insider-outsider theory the more institutions protect the insider, the higher the wage they can achieve (Lindbeck and Snower 1988, Lindbeck and Snower 2002).

2.2 Empirics: Strong evidence for productivity and rent-sharing factors

According to the neo-classical theory, only differences in the labour productivity or in the job-conditions can explain wage differences. Schramm (2004), however, reports that empirical studies on job-conditions (health, ecological, hardness, job-security, working-time and vacations, non-monetary rewards, ...) come to the conclusion that the only robust determinant explaining wage differences is a higher death probability. Thus, dif-

\footnote{The idea is also that workers in smaller firms may be less attainable by trade-unions.}
different wages in the industrial and the service sectors can only be attributed to differences in the ability and productivity of workers in these sectors.

Regarding Europe, Genre, Momferatou, and Mourre (2005) give a broad descriptive analysis of the wage differentials within the Euro Area. Using essentially OECD data, their wage data differ from our data. Nevertheless, they find similar results: the ranking of the countries did not change much over time. On average higher wages are paid in the industry compared to the service sector. This cannot be fully explained by part-time, self-employment or age effects nor by qualification effects (that even have an opposite effect). Gender appears to play no role. The individual productivity-approach however cannot fully explain the differences. The alternative theories provide variables as capital intensity, average firm size and sectoral labour productivity. This means that rent-sharing theories are in line with the observed data.

Haisken-DeNew and Schmidt (1999) confirm with more recent data for Germany and the US that even after controlling for human capital components, job characteristics, status and geographical factors, the inter-industry wage structure is very persistent on both sides of the Atlantic.

Freeman and Schettkat (2001) compare the skill distributions among the unemployed and the employment in USA and Germany, as well as the wage distribution in these two countries and come to the conclusion that the productivity-theory, i.e. the neoclassical one, is not able to explain the better employment performance in the USA.

In most studies focussing on the USA the inter-industry wage differentials are found to be stable over time, even after controlling for unionization and observable characteristics of the workers and job characteristics. Since these differences are highly stable over time, they cannot be attributed to transitory supply or demand shocks affecting a specific industrial sector. Thus, the competitive theories seem to be unable to explain these differentials. Some authors argue that non-standard theories would do a better job; efficiency wage theories for example can explain why the quit rate in high-wage-industries is lower (Krueger and Summers 1988), or rent-sharing theories can explain why profitability seems to be higher in high-wage industries (Dickens and Katz 1986). On the other side, some authors argue that non-observed abilities of workers may affect the inter-industry wage differentials even in the long-run, since these abilities may be judged or perceived differently across industries (Gibbons and Katz 1992). In this latter explanation the competitive model is not challenged anymore, however, the empirical
evidence is not that clear.

Martins, Scarpetta, and Pilat (1996) find that market power exist in the studied industrial sectors and differ across products.

Suedekum and Blien (2007) use a German panel for 326 regional districts and 28 industries & services categories between 1993/5-2002 and find that 70% of the wage differential across regions and industries can be explained by industries and area type fixed effects as well as time-varying characteristics such as qualification (low, medium and high), firm size (small, medium and large), age and gender structure. In a second step they regress the employment growth rate on the unexplained regional wage differential and find out that the elasticity is strongly negative for export-oriented and exposed sectors such as manufacturing and insignificant (and even positive) for rather domestically oriented service sectors (gastronomy and household-related services). Thus, for export-exposed industries, the supply-side effects of wages are dominant whereas for the service sector the demand-side aspects of wages have a bigger impact (however the authors claim that the supply-side effects dominate for both type of industries). It is also remarkable that for Germany the qualification variable is right-signed (branches and regions with higher workers pay higher wages too), the effect of age is slightly positive, as the firm size and the proportion of men.

Erdil and Yetkiner (2001) perform a macroeconomic panel using the STAN-OECD and UNIDO databases. After showing that the wage differentials within the manufacturing sector is constant over the whole available sample period (1970-1992), explanation for these differentials are put forward. For the OECD countries only labour productivity and gender have the right sign and robust effects. Profitability and competitiveness seem to play a role only in 1990. However, their dataset is quite restrictive regarding possible control-variables; working-time, nationality or race, qualification, country size cannot be accounted for example.
3  Bivariate analysis: A correlation assessment

3.1  The data to explain

3.1.1  Sources and definitions

The labour costs analysed in this study are taken from the Labour Cost Statistics (Eurostat)\(^2\).

The wage costs per hour include social security contributions of both employees and employers, as well as the personal income tax paid by the employees. Some additional and quite negligible costs paid by the employers are also included in the statistics (working clothes, ...).

The countries covered are the EU-27 countries; i.e. EU-25+2: Euro Area(12) + UK, Denmark, Sweden and the new 10 Eastern countries + Romania and Bulgaria\(^3\). While Ireland is completely missing, others have very few data (Hungary, Malta and Italy for example).

Note, that for the EU-15 only those employees are considered which are working in firms employing at least 10 employees, whereas for the new member states all are considered. This implies that the wage cost levels are probably over-estimated in the EU-15, since small firms pay lower wages on average (see detailed Eurostat statistics for the NMS-10 grouped by firm size).

In this study we focus on the sectors C to K with CDE = Industry (without construction) and GHIJK = business services (without state and near-state services like health care or education)\(^4\).

The study is based on annual data covering the years 1995 to 2005. For some countries the data base is incomplete, i.e. we face an unbalanced panel.


\(^3\)As the data ends in 2005, Slovenia is not a member of the Euro Area but belongs to the 10 new member states.

\(^4\)More and more data for the other sectors (AB and LMNO) are available. However, these sectors are either negligible in size or cannot be considered as market-driven and are thus not considered in this study.
3.1.2 Two interesting results

1. Contrary to the broadly accepted view (Sinn 2007, Schröder 2007) Germany is not the country with the highest wage costs. In fact, several Northern and Western European countries form a bulk of high-wage countries, while Germany has a rather middle position within the EU-15. The Southern countries and especially the new member states are well below the Western average. See Chart 1 and a detailed descriptive analysis in Dülmann, Hohlfeld, Horn, Logeay, Rietzler, Stephan, and Zwiener (2006) and Horn, Logeay, Stephan, and Zwiener (2007).

2. The second interesting result is the fact that in Germany the wage costs in the industry sector differ largely from those in the service sector. Explaining this puzzling result is the aim of this paper.

Thus, the dependent variable in our empirical study is the hourly wage costs in the industry (CDE) in relation to those in the market service sector (GHIJK). A value above 1 indicates that the average wage costs in the industry are higher than those in
the service sector. Germany has the highest value with 1.25, meaning that industry sectors pay on average 25% more than service sectors! At the opposite the Portuguese wage costs in service sector are on average about 20 to 30% higher than those in the industry\(^5\). As we can see in Chart [2], this result is stable over the available years. A distinction between the “old” European countries (EU-15) and the others (EU-27) is made.

Chart 2: Relative wage costs in Europe: Industry/Services

Source: Eurostat, own calculations. Some countries are missing (Ireland for example).

3.2 Individual factors

As introduced above, we focus on different categories of explanatory variables. The first set which is in line with all theories are the individual features. These variables

---

\(^5\)Interestingly the contribution of domestic demand to GDP growth was 0.8%-points in Germany and 3.7%-points in Portugal between 1995-2005!
are indicators for the individual productivity (stemming from working experience or education), or for some inherent social discrimination (that per se says nothing about the individual productivity but rather how high social discrimination is or may indicate some hidden specific job characteristics). The correlations are shown in Table 1 (p. 14).

3.2.1 Age/Experience

Data for the age composition of the sectoral labour force is available from the Labour Force Survey (Eurostat). Two variables are constructed: a variable for young employees (15-24 years old: $alter_{15-24}$) and one for elder employees (more than 50 years: $alter_{50+}$). The variables are constructed similar to the wage costs: a value above one means that the proportion of younger resp. elder employees in the industry is higher than in the service sector.

$$alter_{15-24} = \frac{\text{Young employees in the industry}}{\text{All employees in the industry}} / \frac{\text{Young employees in the services}}{\text{All employees in the services}}$$

$$alter_{50+} = \frac{\text{Elder employees in the industry}}{\text{All employees in the industry}} / \frac{\text{Elder employees in the services}}{\text{All employees in the services}}$$

The age can be regarded as a proxy for the working experience and thus the acquired qualification of the employee. Usually, a hump-shaped productivity-age-curve is assumed. However, in empirical studies the curve is rather flat at the end, so that a real decline at higher ages is probably not significant. Therefore, a negative sign for the 15-24-variable is expected and a presumably negative albeit not significant correlation for the 50+-variable.

Actually, the correlations are not significant, an indication that either the age plays no role, or it is a rather poor indicator for acquired qualification or of course other counter-acting effects linked to the age are offsetting the presumed effects. However, it is worth noting that from the scatter-plots in the annex (p. 33), Luxembourg seems to follow another logic than the rest of the EU. If we exclude this country, the correlation between $alter_{50+}$ and the relative wage costs is significantly negative for both the EU-15 and EU-27 with ca. -35%. Therefore, the hypothesis of an hump-shaped age-productivity-curve is not true or there are strong institutional features (seniority clauses) of the wage bargaining inducing higher wages for elder employees, independently from their productivity.
3.2.2 Gender

Data for the gender composition of the sectoral labour force is available from the Labour Force Survey (Eurostat). The variable \( \text{gender} \) is constructed similar to the wage costs: a value above one implies that the proportion of men in the industry is higher than the proportion of men in the service sector.

\[
\text{gender} = \frac{\text{Male employees in the industry}}{\text{All employees in the industry}} / \frac{\text{Male employees in the services}}{\text{All employees in the services}}
\] (3)

There is no “theoretical” explanation – along the productivity-theories – why women should earn less than men. But it is a stylized fact reported in various and numerous gender studies, that women are discriminated with respect to wages. Thus, omitting this variable could yield to a severe bias in quantifying the effects of the other variables. This variable is correlated with the part-time proportion, probably indicating a difference in job quality between men and women.

The sign of the expected effect is positive. And indeed the correlation is strong and positive.

3.2.3 Qualification

Data for the formal qualification composition of the sectoral labour force are available from the Labour Force Survey (Eurostat), according to the ISCED classification. The employed persons are divided into three groups according to the highest level of education attained: low, medium and high qualifications. For each sub-sectoral group, a weighted mean was calculated. Again, if the variable \( \text{qualification} \) is above one it means that on average employees in the industry are (formally) better qualified than those in the service sector.

\[
\text{qualification} = \frac{\text{Weighted mean of the qualification categorical variable in the industry}}{\text{Weighted mean of the qualification categorical variable in the services}}
\] (4)

\(^6\)Low: At most lower secondary (ISCED 0-2); Medium: Upper secondary (ISCED 3-4); High: Tertiary (ISCED 5-6). See:
As for the acquired qualification, a positive effect (more qualification should go along with higher wages and thus higher wage costs) is expected here. On the other side, as for the acquired qualification, only the individual qualification is taken into account here. Labour productivity gains induced by using specific technologies or capital inputs in a broad sense may not be well reflected in this variable.

Actually, the correlations are significantly negative, which is in contrast to standard theory. This result was also found by the ECB (Genre, Momferatou, and Mourre 2005). We interpret this result in a way that either this variable is a poor indicator for individual and formal qualification, or that this variable is correlated with some sectoral specific features that destroy the true (positive) presumed correlation. In any case, this result remains puzzling because one would expect that the necessary bias induced by these bivariate considerations should not be so severe as to even produce a counter-intuitive sign.

Table 1: Correlation between individual factors and the wage costs differential between Industry and Services for the EU-27 and the EU-15.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative wage costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALTER_1524 (15-24 years/all employees)</td>
<td>53.5%</td>
<td>80.2%</td>
</tr>
<tr>
<td>ALTER_50+* (50+/all employees)</td>
<td>28.2%</td>
<td>31.8%</td>
</tr>
<tr>
<td>GENDER (Men/all employees)</td>
<td>42.0%</td>
<td>59.0%</td>
</tr>
<tr>
<td>QUALIFICATION (weighted mean: 1-low, 2-mid, 3-high)</td>
<td>-33.5%</td>
<td>-27.9%</td>
</tr>
</tbody>
</table>

Bold figures are significant at the 1%-level. The first number is the correlation coefficient ($\rho$), the second the probability associated with H0: $\rho = 0$ and the last number is the number of observations.

*: w/0 Luxembourg due to big outlier.
3.3 Firm and sectoral factors

The second set of possible explanatory variables are the firm and sector-specific variables. They are also indicators for productivity in the sense that they are in fact not linked to the personal abilities of the workers, but through capital-intensity still affect the overall labour productivity. They can also be interpreted in line with the rent-sharing theories; the bigger the firm, the more likely it possesses some market power and therefore receives a monopolistic rent that can be shared with the workers, leading to wage differences that cannot be explained due to productivity differences. The correlations are shown in Table 2 on p. 17.

3.3.1 Firm size/capital intensity

Data on firm size are very difficult to obtain and are even more difficult to get for a long time period. Data from Eurostat could be collected essentially for the middle of the sample (1999-2001; Eurostat, annual Structural Business Statistics – SBS). Because of a lack of data, some calculations based on the national accounts statistics were performed. Therefore, one should not rely too much on this variable since the measurement error is likely to be important. The variable (\(firmsize\)) is also constructed similar to the wage costs variable. A value above one means that firms in the industry are larger on average than those in service sector.

\[
firmsize = \frac{\text{Average number of employees per firm in industry}}{\text{Average number of employees per firm in services}}
\]  

(5)

This variable is a proxy for capital intensity or the degree of market power in the sector: the larger a firm, the bigger the economies of scale, and thus the larger the amount of capital per head but also the more likely the fact that the firm possesses market power. The first explanation yields a positive effect on labour productivity and hence on wages, the second through the rent-sharing among employers and workers a positive effect on wages too. Thus, in both cases, a positive correlation is expected but cannot be shown for the EU-15. Due to the small number of observations, it is difficult to state significance at all.

For Germany, however, the Federal Statistical Office (Destatis 2006, p. 15, Chart 4) shows that the size of firms and the wage costs are positively correlated and that this
is a long-term feature (1992-2004). Capital intensity (measured as GCF/head) is also positively correlated through the German branches in 2004 with the amount of the wage costs.

3.3.2 Self-employment ratio/capital intensity

Data on self-employment and employment in a sectoral disaggregation can be taken from the national account statistics (Eurostat) but we choose the labour force survey as it is more complete (at least for the EU-15 countries). The variable \( \text{selfemploy} \) takes a value above one if the proportion of self-employed in the industry is larger than those in the service sector.

\[
\text{selfemploy} = \frac{\text{Self-employed/All employment, in the industry}}{\text{Self-employed/All employment, in the services}}
\]  

This variable is also a proxy for the firm size (the correlation between the two variables is indeed significantly negative with about -65%) and consequently for the capital intensity and/or for the market power of the firms. Because the link is reversed (a larger proportion of self-employed goes along with a smaller average firm size), we expect a negative sign. Indeed the correlation is negative, however it is not significant.

3.3.3 Part-time

Data for part-time jobs on a sectoral level are available from Eurostat only for the year 2000 and from the labour costs statistics. Thus this variable is quite restricted. Again the variable takes values above one if the part-time ratio is higher in the industry than in the service sector.\(^7\)

\[
\text{teilzeit} = \frac{\text{Part-time employees/All employees, in the industry}}{\text{Part-time employees/All employees, in the services}}
\]

Similar to gender, there is no “theoretical” justification for an influence of this variable, as the labor costs are already corrected for the hours worked. However, like for gender, it is a stylized fact that the hourly wages for part-time employees are lower than

\(^7\)This was calculated for employees of firms which employ at least ten employees.
those of full-time employees (OECD 1999, p. 22-25). Thus, one should take this into account. This variable is negatively correlated with the gender variable (-30%, meaning that if the fraction of female employees is high, part-time jobs are high too) as mentioned above, so that this variable may be an indicator for social discrimination and for job quality. This would yield a negative correlation.

The correlation is indeed negative for the EU-27 (albeit not significant) but is unexpectedly positive for the EU-15 (but not significant either). These correlations should be interpreted with caution because of the very small sample size.

3.3.4 (Labour) Productivity

Productivity measures are only available on a per-capita basis, because figures for hours are not available for all countries. Working time reductions have reduced the productivity per head but increased wages per hour in the past. Besides, taking the ratio of sectoral gross value added to the sectoral employment levels is thought to be highly endogenous and may yield to strong endogeneity problems. Therefore, productivity per capita cannot be used here.

Table 2: Correlation between firm-specific factors and the wage costs differential between Industry and Services for the EU-27 and the EU-15.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FIRMSIZE</td>
<td>12.0%</td>
<td>-13.1%</td>
</tr>
<tr>
<td></td>
<td>31.9%</td>
<td>42.8%</td>
</tr>
<tr>
<td></td>
<td>71</td>
<td>39</td>
</tr>
<tr>
<td>SELFEMPLOY*</td>
<td>-15.8%</td>
<td>-33.9%</td>
</tr>
<tr>
<td>(Selfemployed/all employment)</td>
<td>2.3%</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>207</td>
<td>175</td>
</tr>
<tr>
<td>TEILZEIT</td>
<td>-12.1%</td>
<td>15.9%</td>
</tr>
<tr>
<td>(Part-time/all employees)</td>
<td>57.3%</td>
<td>58.7%</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>14</td>
</tr>
</tbody>
</table>

Bold figures are significant at the 1%-level. The first number is the correlation coefficient ($\rho$), the second the probability associated with H0: $\rho = 0$ and the last Number is the number of observations.

*: w/o Luxembourg due to big outlier.
3.4 Country specific factors: institutional settings

The third set of possible explanatory variables are the country-specific variables. Clearly, they are indicators for failures of the neo-classical framework, as they indicate some institutional features of the labour market and show whether market power plays a significant role. The correlations are shown in Table [3].

3.4.1 Unemployment rate/strength of workers

Data on the standardized unemployment rate are provided by Eurostat (ILO-concept). The unemployment rate here is the national-wide rate (no sectoral differences are available). This should be an indicator for the market power or relation between employers and employees as predicted by the insider/outsider theory: A higher unemployment rate goes along with a weaker bargaining position of employees. By which extent this should affect employees in industrial sectors more than in service sectors is an open question. But if no other indicators are included, as in the years observed (1995-2004), the structural change operates in favor of services: A higher unemployment rate should weaken the industry employees more than those in the service sectors on the one hand (negative effect), on the other hand a higher unemployment rate may be rather an indication for poor macroeconomic performance, which could weaken the service sector more than the industry (positive effect).

The correlation is significantly positive which gives more evidence for the second explanation.

3.4.2 Minimum wage/institutional factors

A dummy variable takes value 1 if the country has a minimum wage and 0 otherwise. There are only seven countries without legal national minimum wage throughout the sample (Austria, Germany, Finland, Italy, Denmark, Sweden and Cyprus), the UK has

\[\text{Industry-firms are thought to be more likely to export their products than Services in general. This is confirmed by the Input-Output tables for Belgium, Germany, Ireland, Spain, France, Italy, Netherlands, Austria, Portugal, Finland, Denmark, Sweden, UK, Lithuania, Hungary, Slovenia and Slovakia for the available years between 1995 and 2004. By this way, industrial firms may compensate domestic slackness through better foreign demand more significantly than firms in the Services.}\]
adopted a minimum wage in 1999.

Another interesting variable would be the proportion of employees covered by minimum wage(s); as we exclude civil servants and employees of the social services, for countries where the minimum wage is legal and national, this coverage rate will be 100%. More interesting is the proportion covered in the North-South axis of countries that do have minimum wages only at sectoral levels and set by collective agreements (Italy, Germany, Austria, Denmark and the three scandinavian countries). From Funk and Lesch (2005) and Husson (2006) some figures are available for some years. It appears that Austria, Denmark, Finland and Italy have high coverage rates (above 85%) whereas Germany and Norway have low coverage rates (around 70% in 2004, but here we suspect that this concerns only West-Germany\(^9\)). A closer look at low-pay sectors that are all in the Services (textile/clothing, retail, hotel/restaurants, hairdressing) also shows that the coverage of the minimum wages set by the collective agreements are much smaller in those two countries than in the others. Due to lack of data however, the aim to construct such a variable was abandoned.

A national-wide minimum wage affects (almost) all employees. It should reduce the wage differentials because it increases the very low wages in all sectors. Thus the absence of a minimum wage should be associated with more extreme values of our dependent variable and the presence of a national minimum wage with values concentrated around one. This non-linear effect can be caught if we divide the sample in two (countries/years with a value of the dependent below one and countries/years with a value above one). In the first group a positive coefficient is expected and in the second a negative one.

For the whole sample, a significant negative correlation for both variables can be observed: the presence of a minimum wage is significant. The higher its level, the lower the industry/services wage gap. For the split sample, the expected non-linear effect is not met. Even counter-intuitive, the negative correlation in the first sub-sample appears to be twice to three times higher than in the second sub-sample (see Table 4).

\(^9\)In Belgium the minimum wage concerns only the employees in the private sector. As we only look at the private sector, excluding construction and the primary sectors, the coverage rate is 100%. In Cyprus, the minimum wage concerns only some professions; sales staff, clerical workers, auxiliary health care staff and auxiliary staff in nursery schools, crèches and schools. Thus the coverage rate is not 100%.

\(^{10}\)From the WSI-Tarifarchiv (2006) figures, the coverage rate by collective agreements, where minimum wages are normally defined, varied from 76% to 67% between 1998 and 2005 in West-Germany and from 63% to 53% in the East.
3.4.3 Union density rate/institutional factors

The union density rate is taken from the OECD-database (Bassanini and Duval 2006). The data set ends in 2003 and covers only Western European countries.

One would expect that the union density is especially low in service sector, as mentioned above. Thus, a higher union density would rather go along with a smaller difference between union coverages in the service compared to the industry sector. This would support a negative effect in the labour cost relation between Industry and Services. The correlation here is negative but not significant.

It is remarkable that a non-linear effect can be found (the correlation coefficient is significant positive for values of the dependent below one and significant negative for values of the dependent above one; See Table 4). The explanation is quite straightforward and follows the arguments for a hump-shaped wage/centralization degree of wage bargaining à la Calmfors and Driffill (1988). The stronger the trade-unions, the more likely they coordinate their actions and promote wage compression over all sectors leading to a reduction of the wage differential. On the other hand, the less representative, the more firm-specific the trade-unions demands will be, yielding a higher wage dispersion across sectors. This means that values well above and below one in our dependent variable should be associated with low values of union density, and values about one of our dependent variable should be associated with high values of union density. This means that the correlation should be positive for values of the dependent below one and negative for values above one.

3.5 Growth composition indicators

The idea for this fourth and last set of variables is that the demand addressed to a sector plays at least as an important role in the wage determination as the relative competitiveness of this sector. Thus, we think that the wage costs in a sector are not only reflecting the (marginal) productivity of this sector net of the employer’s rent due to imperfect competition but also the structure of the demand (foreign versus domestic demand) a specific sector is confronted with.

Here, however, a precision has to be made. As the sectoral wage differentials are persistent over time i.e. over the business cycles, the state of demand, reflected within
Table 3: Correlation between country-specific factors and the wage costs differential between industry and services for the EU-25 and the EU-15.

<table>
<thead>
<tr>
<th></th>
<th>EU-27</th>
<th>EU-15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1995-2005</td>
<td></td>
</tr>
<tr>
<td>Relative wage costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALQ</td>
<td>14.4%</td>
<td>27.9%</td>
</tr>
<tr>
<td>(Unemployment rate)</td>
<td>3.0%</td>
<td>0.1%</td>
</tr>
<tr>
<td></td>
<td>226</td>
<td>135</td>
</tr>
<tr>
<td>MINWAGE_IS</td>
<td>-20.1%</td>
<td>-15.7%</td>
</tr>
<tr>
<td>(0: no minimum wage, 1: national MW)</td>
<td>0.2%</td>
<td>6.9%</td>
</tr>
<tr>
<td></td>
<td>242</td>
<td>135</td>
</tr>
<tr>
<td>UDENS</td>
<td></td>
<td>-13.5%</td>
</tr>
<tr>
<td>(Union Density)</td>
<td></td>
<td>19.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95</td>
</tr>
</tbody>
</table>

Bold figures are significant at the 1%-level. The first number is the correlation coefficient (ρ), the second the probability associated with H0: ρ = 0 and the last number is the number of observations.

Table 4: Non-linear correlation between country-specific factors and the wage costs differential between Industry and Services for the EU-25 and the EU-15.

<table>
<thead>
<tr>
<th></th>
<th>EU-27</th>
<th>EU-15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w &lt; 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w &gt; 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALQ</td>
<td>14.4%</td>
<td>28.5%</td>
</tr>
<tr>
<td></td>
<td>3.0%</td>
<td>0.2%</td>
</tr>
<tr>
<td></td>
<td>226</td>
<td>114</td>
</tr>
<tr>
<td>MINWAGE_IS</td>
<td>-20.1%</td>
<td>-41.7%</td>
</tr>
<tr>
<td></td>
<td>0.2%</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>242</td>
<td>122</td>
</tr>
<tr>
<td>UDENS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-13.5%</td>
<td>66.3%</td>
</tr>
<tr>
<td></td>
<td>19.2%</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>33</td>
</tr>
</tbody>
</table>

Bold figures are significant at the 1%-level. The first number is the correlation coefficient (ρ), the second the probability associated with H0: ρ = 0 and the last number is the number of observations.

traditional business cycle indicators like the growth rate, cannot explain the wage gap. Here, we think rather in terms of long-term structure or strategies; Germany for example has a long tradition of using external trade to foster growth, whereas other European countries push rather domestic demand. In that respect, it is likely that the participation in the Monetary Union has induced a change in preferences in economic policy in those
countries which traditionally focus rather on domestic forces than in countries with a long experience of de facto monetary bindings like Germany, Netherlands or Austria.

In these respects, we thought that the indicator should be cumulative and relative, i.e. has a chance to persist over-business cycles, and reflects long-lasting differences in the demand addressed to the industry and service sector. We look at the composition of growth as an explanatory factor for the wage differential; industry sectors are traditionally more export-oriented, whereas the service sector typically depends on the domestic market (with some exceptions of course). Growth will not have the same impact on the wage opportunities for workers in the industry and the service sector, depending on the driving forces; the German model of export-led growth should favour the industry more than the service sector whereas the Portuguese model of domestic demand-led growth should be more favourable for the services. The correlations are shown in Table 5.

Along this idea, cumulative and relative growth variables for the exports and domestic demand are constructed: the growth rate in % of exports resp. of domestic demand resp. of GDP was cumulated from 1996 onwards, calculating an index with value 100 in 1995 (Malta and Romania start only 1999 due to lack of data); This index is then put in relation to the one of the Euro Area. A value above 100 means therefore that the accumulated growth performance since 1995 is above average.

\[
\text{growth}_{name}^{i,t} = \frac{\prod_{t=1996}^{t} \left( \frac{name_{i,t}}{name_{i,t-1}} \right)}{\prod_{t=1996}^{t} \left( \frac{EUR_{12,t}}{EUR_{12,t-1}} \right)} \times 100
\]  

(8)

With name=exp, gdp or intdd for resp. exports, GDP and domestic demand in constant terms; i the country index and t the time index. The index starts in 1995 with the value 100.

We calculate this indicator for exports, domestic demand and GDP at constant prices. The expected effects are positive for the first one, negative for the second one; for the third one it depends on the fact whether the first or the second effect is dominant. The correlations have the expected signs and are significant for the domestic demand and GDP. The results show that development of the domestic demand is much more important than export demand.
Table 5: Correlation between growth structure indicators factors and the wage-cost differential between Industry and Services for the EU-25 and the EU-15.

<table>
<thead>
<tr>
<th></th>
<th>EU-27</th>
<th>EU-15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1995-2005</td>
<td></td>
</tr>
<tr>
<td>Relative wage costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GROWTH_EXPK2</td>
<td>-1.3%</td>
<td>22.5%</td>
</tr>
<tr>
<td>(cumulated export-growth relative to EUR-12)</td>
<td>84.3%</td>
<td>0.9%</td>
</tr>
<tr>
<td></td>
<td>242</td>
<td>135</td>
</tr>
<tr>
<td>GROWTH_GDPK2</td>
<td>-24.6%</td>
<td>-29.7%</td>
</tr>
<tr>
<td>(cumulated GDP-growth relative to EUR-12)</td>
<td>0.0%</td>
<td>0.1%</td>
</tr>
<tr>
<td></td>
<td>242</td>
<td>135</td>
</tr>
<tr>
<td>GROWTH_INTDDK2</td>
<td>-23.7%</td>
<td>-41.1%</td>
</tr>
<tr>
<td>(cumulated domestic demand-growth relative to EUR-12)</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>242</td>
<td>135</td>
</tr>
</tbody>
</table>

Bold figures are significant at the 1%-level. The first number is the correlation coefficient ($\rho$), the second the probability associated with H0: $\rho = 0$ and the last number is the number of observations.

4 Multivariate analysis

The advantage of performing a multivariate analysis which includes both dimensions (time and cross-sections) is twofold: first with an increasing number of observations, we expect to obtain more accurate estimates for the effects of individual variables on the relative wage differential and second – contrary to the bivariate correlations – in a multivariate analysis, one controls for different effects at the same time and should be able to isolate the genuine effect of a factor, reducing therefore potential omitted variable bias.

The regressors we use, summarized in $X_{it}$, are [$alter_{1524}$, $alter_{50+}$, gender, qualification, selfemploy, alq, minwage_is, growth_expk2, growth_intdd2k]. We dropped firmsize, udens and teilzeit because these variables had too few data. Still the data set remains unbalanced. All variables are time-variable, whereas the dummy minwage_is actually is time-invariant for all countries but for the UK\[11\].

\[11\]It takes the value 0 between 1995 and 1998 and 1 after.
4.1 Static regressions

It is not obvious that the model should contain a dynamic term, since the series are quite stable (they are all relative and thus do not have trends). Because our data set is a macro-panel with countries from the EU-27, it seems also reasonable to rule out random-effects. At least a discussion about which variable may be correlated with country-specific unobserved characteristics should be done.

In principle, all variables may be correlated with the country-effects. But certainly the unemployment rate, the presence or not of a minimum wage may be very likely candidates. As we have only the minimum-wage dummy that may be considered as almost time-invariant, the fixed-effects estimator may be the most appropriate.

\[ lohnkosten_{it} = X_{it}' \beta + \lambda' D T \_t + \alpha_0 + \alpha_i + u_{it} \] (9)

The results of the standard estimators are reported in Table 6. The Hausman-test (H0: Random Effects are present) is rejected and the Breusch-Pagan test (H0: Fixed effects are present) is consistently not rejected. This is in line with our intuition that a fixed-effect model would better describe a macro-panel. The pooled estimation is reported below:

<table>
<thead>
<tr>
<th></th>
<th>coefficient</th>
<th>H-robust p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALTER_1524</td>
<td>-4.58E-02</td>
<td>60.2%</td>
</tr>
<tr>
<td>ALTER_50PLUS</td>
<td>-2.51E-03</td>
<td>98.3%</td>
</tr>
<tr>
<td>GENDER</td>
<td>3.88E-01</td>
<td>1.5%</td>
</tr>
<tr>
<td>QUALIFICATION</td>
<td>-4.70E-02</td>
<td>11.2%</td>
</tr>
<tr>
<td>SELFEMPLOY</td>
<td>-3.21E-02</td>
<td>83.7%</td>
</tr>
<tr>
<td>ALQ</td>
<td>1.18E-02</td>
<td>3.7%</td>
</tr>
<tr>
<td>MINWAGE_IS</td>
<td>-2.60E-03</td>
<td>94.4%</td>
</tr>
<tr>
<td>GROWTH_EXPK2</td>
<td>5.70E-04</td>
<td>45.4%</td>
</tr>
<tr>
<td>GROWTH_INTDDK2</td>
<td>-1.46E-03</td>
<td>19.4%</td>
</tr>
</tbody>
</table>

Wald (X): $\chi^2(9) = 34.55 [0.0\%]$ ; Wald (time): $\chi^2(10) = 16.94 [7.6\%]$ AR-1 & 2 rejected at 1%, with PcGive.
Table 6: Static regressions: results of the standard estimators

<table>
<thead>
<tr>
<th>Estimation method:</th>
<th>xtreg</th>
<th>xtgee</th>
<th>xtgls</th>
<th>xtpcse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fixed</td>
<td>Random</td>
<td>Random</td>
<td>Random</td>
</tr>
<tr>
<td></td>
<td>effects</td>
<td>GLS</td>
<td>MLE</td>
<td>GLS</td>
</tr>
<tr>
<td></td>
<td>(within)</td>
<td>GLS</td>
<td>MLE</td>
<td>GLS</td>
</tr>
<tr>
<td>Alter_1524</td>
<td>38.6E-3</td>
<td>33.7E-3</td>
<td>33.5E-3</td>
<td>33.7E-3</td>
</tr>
<tr>
<td>Alter_50plus</td>
<td>43.2E-3</td>
<td>39.2E-3</td>
<td>39.0E-3</td>
<td>39.1E-3</td>
</tr>
<tr>
<td>Gender</td>
<td>80.9E-3</td>
<td>112.4E-3</td>
<td>113.6E-3</td>
<td>112.9E-3</td>
</tr>
<tr>
<td>Qualification</td>
<td>-3.1E-3</td>
<td>-9.7E-3</td>
<td>-9.9E-3</td>
<td>-9.8E-3</td>
</tr>
<tr>
<td>Selfemploy</td>
<td>5.7E-3</td>
<td>-8.2E-3</td>
<td>-8.7E-3</td>
<td>-8.4E-3</td>
</tr>
<tr>
<td>Alq</td>
<td>-3.6E-3</td>
<td>-2.0E-3</td>
<td>-1.9E-3</td>
<td>-2.0E-3</td>
</tr>
<tr>
<td>Minwage_is</td>
<td>-8.6E-3</td>
<td>-9.4E-3</td>
<td>-9.4E-3</td>
<td>-9.4E-3</td>
</tr>
<tr>
<td>Growth_explk2</td>
<td>177.7E-6</td>
<td>212.1E-6</td>
<td>213.3E-6</td>
<td>212.5E-6</td>
</tr>
<tr>
<td>Growth_intld2</td>
<td>-1.9E-3</td>
<td>-1.7E-3</td>
<td>-1.7E-3</td>
<td>-1.7E-3</td>
</tr>
<tr>
<td>Country effect</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Time effect</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>nb.obs.</td>
<td>201</td>
<td>201</td>
<td>201</td>
<td>201</td>
</tr>
<tr>
<td>nb.param.</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>R² within</td>
<td>19.4%</td>
<td>19.0%</td>
<td>19.0%</td>
<td>19.0%</td>
</tr>
<tr>
<td>R² between</td>
<td>3.5%</td>
<td>8.5%</td>
<td>8.6%</td>
<td>8.6%</td>
</tr>
<tr>
<td>R² overall</td>
<td>6.7%</td>
<td>13.6%</td>
<td>13.7%</td>
<td>13.7%</td>
</tr>
<tr>
<td>Prob&gt;F</td>
<td>1.2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob&gt;F (TD=0)</td>
<td>25.3%</td>
<td>0.0%</td>
<td>32.2%</td>
<td>39.6%</td>
</tr>
</tbody>
</table>

xtreg: OLS estimator; xtgee: ML-estimator as in xtreg but handled unbalanced panel differently; xtgls: GLS-estimator with assumed heteroskedasticity but no correlation across panels and a ar(1)-structure across time; xtpcse: GLS-estimator with correlation in both dimensions (time and cross-sections) and heteroskedasticity across panels; psar: the assumed correlation is decreasing with the time distance.

*: with small-sample correction for the standard errors (Swamy-Arora) and non-constant theta.
**,**,**: significant at 10%, 5%, 1% level.
5 Conclusions

In this article we analyze the new data provided by Eurostat on hourly wage costs, which are comparable across all European countries. Among other interesting features, a surprising result is the outlier position of Germany where in the industry sector wages are on average 25% higher than in the service sector. The aim of this paper is to explain this puzzle.

In a first and descriptive step, a bivariate correlation analysis was performed and the following results were found: only gender is a significant factor at the individual level. Especially the (formal) qualification shows significant wrong-signed correlation, a contradiction with the widely asserted low productivity in services. On the contrary, firm specific factors indicating the importance of capital intensity in explaining wage differentials (self-employment share) do not seem to have a great explanatory power. Factors (country specific and demand indicators) that are in accordance with non-standard wage determination theories are found significant (unemployment rate, minimum wage, growth structure).

In a second step a multivariate analysis is performed, allowing to control for each effect. The signs found in bivariate correlations are so far confirmed for the elder-age-variable (+), gender (+), qualification (counter-intuitive -) and the domestic demand growth variable (-).

From these preliminary results, it seems that our hypothesis is confirmed; the structure of the demand addressed to the sectors is quite an important factor explaining the development of wages and therefore the EU sectoral wage differentials.
References


6 Annex: Overview of the estimators used

6.1 Pooled OLS

In the pooled model, there is no country effects and the time effects are common to all individuals. The errors are supposed to be uncorrelated in the time dimension and across countries and homoskedastic in the cross-sectional dimension:

\[ y_{it} = \alpha + X_{it}'\beta + u_{it} \] (10)

In STATA, the command is xtreg with time dummies.

6.2 Static Fixed Effects (SFE)

In the fixed effects model, the error term is decomposed in a fixed unit-specific component (a country-specific intercept) and an observational error term. There exists several methods to estimate such a model; Either one adds individual dummies (LSDV-estimator):

\[ y_{it} = \alpha'DI_i + \lambda'DT_t + X'_{it}\beta + u_{it} \] (11)
Or one can transform the variables (subtract the group mean), so that the time-invariant variables are dropped from the model (WITHIN-estimator):

\[ y_{it} - \bar{y}_i = (\alpha_i - \alpha_i) + X'(DT_t - \bar{DT}) + (X_{it} - \bar{X}_i)'\beta + (u_{it} - \bar{u}_i) \quad (12) \]

This is achieved in Stata with the command `xtreg` and the option “fe”. In both cases (that yield exactly the same estimates for the \( \beta \), as the OLS method is applied, we need to have homoskedastic and uncorrelated errors, and that they are unrelated to the \( \alpha_i \) and the other exogenous factors, for the estimators to be consistent and BLUE.

### 6.3 Static Random Effects (SRE)

In the fixed effects model, the error term is decomposed in a random unit-specific component (a country-specific error term) and an observational error term. As now it is clear that the overall error term will not met the assumption of autocorrelation and probably also not of homoskedasticity, the GLS-estimator is required. One can show further that the GLS reduce to OLS applied to the following transformed model:

\[ y_{it} - \theta_i y_i = (1 - \theta_i)\alpha^* + X'(DT_t - \theta_i\bar{DT}) + (X_{it} - \theta_i\bar{X}_i)'\beta + \varepsilon_{it} - \theta_i\varepsilon_{it} \quad (13) \]

\( \theta_i \) stands for \( 1 - \sqrt{\frac{\sigma_u^2}{\sigma_u^2 + T_i\sigma_\alpha^2}} \) and \( \varepsilon_{it} = \alpha_i + u_{it} \).

We need to estimate \( \theta_i \) and there exists therefore different estimators , as we can only perform F-GLS.

In Stata the commands `[xtreg, re]` and `[xtreg, re sa]` call the GLS estimator. In both cases the Swamy-Arora variance estimator based on the within and between regressions is implemented. The second specification has a small sample correction, that differ from the first one in unbalanced panels. Important to note here, is that the \( u_{it} \)-error term possess the usual properties: homoskedastic across individual, uncorrelated between the individuals and in the time-dimension. It is also unrelated to the individual-specific error \( \alpha_i \) and to the other exogenous variables. As this is a F-GLS estimation, test-distributions hold only asymptotically (\( \chi^2 \) rather than F).

It is possible to estimate the F-GLS also with maximum-likelihood `[xtreg, mle]`. In this case, the residual are assumed also to be normally distributed and an iterated regression is performed. In case the total number of observation (here 201) is smaller
than 300 and the data unbalanced (our case too) the [mle] and [re] regressions will yield different results.

If the correlation structure of \( u_i \) is not as simple as assumed above esp. if some time-autocorrelation is present, then the GLS estimator needs to be adapted. There is several ways to take account of richer variance structure.

[xtgee] is the first possibility embraced in the estimation part. It fits a population-averaged panel-data model; the option \( [f(gauss) i(id)] \) consider that the errors are normally distributed and the model linear in the coefficients. This estimator consider then the within-group correlation structure \( R_i \):

\[
\begin{pmatrix}
\text{corr}(u_{i1}, u_{i1}) & \text{corr}(u_{i1}, u_{i2}) & \cdots & \text{corr}(u_{i1}, u_{iT_i}) \\
\text{corr}(u_{i2}, u_{i1}) & \text{corr}(u_{i2}, u_{i2}) & \cdots & \text{corr}(u_{i2}, u_{iT_i}) \\
\vdots & \vdots & \ddots & \vdots \\
\text{corr}(u_{iT_i}, u_{i1}) & \text{corr}(u_{iT_i}, u_{i2}) & \cdots & \text{corr}(u_{iT_i}, u_{iT_i})
\end{pmatrix} =
\begin{pmatrix}
r_{11} & r_{12} & \cdots & r_{1T_i} \\
r_{21} & r_{22} & \cdots & r_{2T_i} \\
\vdots & \vdots & \ddots & \vdots \\
r_{T_i1} & r_{T_i2} & \cdots & r_{T_iT_i}
\end{pmatrix}
\]

Then the within-group correlation structure \( R_i \) can take two forms:

- corr(exch): \( r_{ts} = \begin{cases} 
1 & \text{if } t = s \\
\rho & \text{if } t \neq s
\end{cases} \). This equivalent to the [xtreg, re] and [xtreg, mle] commands if the data are balanced, with [xtreg, pa] otherwise.

- corr(ar1): \( r_{ts} = \begin{cases} 
1 & \text{if } t = s \\
\rho^{|t-s|} & \text{if } t \neq s
\end{cases} \). This equivalent to the [xtreg, re] command.

[xtgls] is another possibility. It is a F-GLS estimator which allows for ar(1)-autocorrelation within panel, as [xtgee], and cross-sectional correlation and cross-sectional heteroskedasticity. The variance of the overall residuals \( (\alpha_i + u_{it}) \) can be written as the Kronecker product of the within-group variance \( (\sigma) \) and the cross-sectional variance \( (\Omega) \):

\[
E(\varepsilon \varepsilon') = \sigma \otimes \Omega = 
\begin{pmatrix}
\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\
\sigma_{21} & \sigma_{22} & \cdots & \sigma_{12N} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{N1} & \sigma_{N2} & \cdots & \sigma_{NN}
\end{pmatrix} \otimes 
\begin{pmatrix}
\Omega_{11} & \Omega_{12} & \cdots & \Omega_{1N} \\
\Omega_{21} & \Omega_{22} & \cdots & \Omega_{12N} \\
\vdots & \vdots & \ddots & \vdots \\
\Omega_{N1} & \Omega_{N2} & \cdots & \Omega_{NN}
\end{pmatrix}
\]

If the diagonal of the \( \sigma \)-matrix is filled with different numbers, the individuals do not
have the same overall variance (panel-heteroskedasticity). If the off-diagonals are not zero, the individuals are correlated with each other (panel-correlation). If the Ω-matrix is not identity, then the residuals are autocorrelated in the time dimension (within-group correlation). With this estimator several options are possible:

- **panels(hetero)**: state that the Ω-matrix is identity (default if no corr-option specified) and $\sigma = \begin{pmatrix} \sigma_{11} & 0 & \ldots & 0 \\ 0 & \sigma_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_{NN} \end{pmatrix}$

- **corr(ar1)**: state that the Ω-matrix is such that the serial correlation structure is $r_{ts} = \begin{cases} 1 & \text{if } t = s \\ \rho_{|t-s|} & \text{if } t \neq s \end{cases}$

- **corr(psar1)**: state that this correlation factor can be different for each individual: $r_{ts,i} = \begin{cases} 1 & \text{if } t = s \\ \rho_{i|t-s|} & \text{if } t \neq s \end{cases}$

The last estimator used in this paper is [xtpcse]; it calculates a panel-corrected standard errors for the OLS-estimates (actually Prais-estimator that corrects for first order autoregression in the time-dimension, i.e. a GLS estimator!). We use the [pairwise] option that specifies that all information available for an individual should be used to calculate the covariance matrix. This has influence only when the data are unbalanced. The option [hetonly] is also selected implying that panel-heteroskedasticity is allowed (see above). Two sorts of within-group autocorrelation is then allowed as above, a common ar(1)-structure (ar1) or individual-specific one (psar1). [xtgls] and [xtpcse] are consistent, and under the correct assumption of the error structure F-GLS would be more efficient. But it is argued that with the typical small samples used in social science, xtpcse should be preferred, as it is more conservative.

7 Annex: Scatter diagrams of the data
Chart 3: Scatter diagram: Age (15-24) vs Wages, EU-27 and EU-15

Chart 4: Scatter diagram: Age (50+) vs Wages, EU-27 and EU-15

Chart 5: Scatter diagram: Gender vs Wages, EU-27 and EU-15

lk_neu: relative hourly wages (industry/services)
Chart 6: Scatter diagram: Qualification vs Wages, EU-27 and EU-15
Chart 7: Scatter diagram: Firm size vs Wages, EU-27 and EU-15

Chart 8: Scatter diagram: Self-employment vs Wages, EU-27 and EU-15

Chart 9: Scatter diagram: Part-time employment vs Wages, EU-27 and EU-15
Chart 10: Scatter diagram: Unemployment rate vs Wages, EU-27 and EU-15

Chart 11: Scatter diagram: Minimum wage-dummy vs Wages, EU-27 and EU-15

Chart 12: Scatter diagram: Union density vs Wages, EU-27 and EU-15
Chart 13: Scatter diagram: Relative export growth vs Wages, EU-27 and EU-15

Chart 14: Scatter diagram: Relative GDP growth vs Wages, EU-27 and EU-15

Chart 15: Scatter diagram: Relative internal demand growth vs Wages, EU-27 and EU-15