# IAW DISCUSSION PAPERS 

# An Almost Ideal Wage Database Harmonizing the ILO October Inquiry 

Daniela Harsch<br>Jörn Kleinert

Institut für Angewandte Wirtschaftsforschung e.V.
Ob dem Himmelreich 1 | 72074 Tübingen | Germany
Tel.: +49707198960 | Fax: +497071989699
ISSN: 1617-5654

## IAW-Diskussionspapiere

Das Institut für Angewandte Wirtschaftsforschung (IAW) Tübingen ist ein unabhängiges außeruniversitäres Forschungsinstitut, das am 17. Juli 1957 auf Initiative von Professor Dr. Hans Peter gegründet wurde. Es hat die Aufgabe, Forschungsergebnisse aus dem Gebiet der Wirtschafts- und Sozialwissenschaften auf Fragen der Wirtschaft anzuwenden. Die Tätigkeit des Instituts konzentriert sich auf empirische Wirtschaftsforschung und Politikberatung.

Dieses IAW-Diskussionspapier können Sie auch von unserer IAW-Homepage als pdf-Datei herunterladen:
http://www.iaw.edu/Publikationen/IAW-Diskussionspapiere

## ISSN 1617-5654

Weitere Publikationen des IAW:

- IAW-News (erscheinen $4 x$ jährlich)
- IAW-Forschungsberichte

Möchten Sie regelmäßig eine unserer Publikationen erhalten, dann wenden Sie sich bitte an uns:

IAW Tübingen, Ob dem Himmelreich 1, 72074 Tübingen, Telefon 07071 / 98 96-0
Fax 07071/9896-99
E-Mail: iaw@iaw.edu

Aktuelle Informationen finden Sie auch im Internet unter:

> http://www.iaw.edu

Der Inhalt der Beiträge in den IAW-Diskussionspapieren liegt in alleiniger Verantwortung der Autorinnen und Autoren und stellt nicht notwendigerweise die Meinung des IAW dar.

# An Almost Ideal Wage Database Harmonizing the ILO October Inquiry 

Daniela Harsch ${ }^{\text {a }}$ Jörn Kleinert ${ }^{\text {b }}$,<br>${ }^{\text {a }}$ University of Tübingen, Mohlstr. 36, 72074 Tübingen, Germany, daniela.harsch@uni-tuebingen.de<br>${ }^{\mathrm{b}}$ University of Graz, Universitätsstr. 15/F4, 8010 Graz, Austria, joern.kleinert@uni-graz.at


#### Abstract

The lack of comprehensive, international comparable wage data has been deplored for decades and has constrained the empirical analysis of wage growth and inequality. This is the case although, since 1924, the International Labor Organization (ILO) has conducted an October Inquiry to obtain data on wages worldwide, which leads to an annual wage survey containing data for 161 occupations in over 130 countries. Freeman and Oostendorp (2000) have started a novel project to make use of the October Inquiry, which we update. We provide a documentation about the several steps taken to transfer the data into a comparable and usable form. The data allow analyzing wage growth and inequality in a comprehensive way. In this paper, we describe the way we converted the data and present some results on developments in the wage structure between and within countries and occupations.


Key words: wage growth, cross section models, wage structure JEL: C31, J31, E24

[^0]
## 1. Introduction

The lack of comprehensive, international comparable wage data has been deplored for a while and has made the analysis of wage growth and inequality for a larger sample of countries hardly possible. Freeman and Oostendorp (2000) have started a great project of wage harmonization that makes use of the October Inquiry database of the International Labor Organization (ILO). They made this rather unused data available for a wider group of researchers by cleaning, correcting, and normalizing the data in order to make the observations comparable across countries and occupations. Unfortunately, the data is still not widely used. Therefore, we try a new start in preparing the data closely following the procedure of Freeman and Oostendorp (2000, 2001).

We prepare four STATA datasets, and a documentation about the several steps taken to transfer the data into a comparable and usable form. We provide different datasets because we believe that there is not one dataset suitable for the different questions researchers have with respect to wage data. The first dataset we provide is a more or less raw dataset, corrected only for structural displacements and typos. We then correct the raw data for labeling mistakes and account for country-specific single events such as a currency reform. We describe the corrections to give everyone the chance to see the changes we made, to correct us (and therefore the data), and to inform us about specific events not yet captured. The corrected and cleaned dataset is the second one we provide. Third, we use the corrected data to construct a dataset that reports standardized wages for every combination of country, year and occupation. Following Freeman and Oostendorp (2000), we chose men's average monthly wage as standard. We provide a standardized dataset that is reduced in observations by reporting only one wage for each country-year-occupation combination. In a fourth step, we enlarge the dataset by reducing the numerous missing observations. For this purpose, we impute a predicted value from a linear prediction. This greatly increases international comparability of the data because the wage data of many countries shows gaps over time and across occupations.

In this paper, we describe how we have transformed the ILO October Inquiry into a consistent database and give a short overview over the wage pattern around the world. In Section 2, we shortly introduce the October Inquiry Data. In Section 3, we describe the challenges posed by the database and illustrate the correction pro-
cedure and the standardization process. In Section 4, we discuss international wage patterns and their evolution over time using the results from the standardization procedure. Section 5 gives some descriptive statistics of the standardized data. In Section 6, we sketch the data imputation. Section 7 summarizes the work and gives an outlook of future work with the data.

## 2. Data

Since 1924, the International Labor Organization (ILO) has conducted an October Inquiry to obtain data on wages and hours worked for a large number of countries and occupations all over the world. Every year, the ILO sends questionnaires to national governments asking for detailed information about wages, hours of work, and occupations. This leads to an annual wage survey which contains data covering up to $161^{1}$ occupations in 49 industries for more than 130 countries. Since there are large gaps in the data, it is only usable from the beginning of the 1980s on, although a larger period of time is available. For our analysis, we choose the time period from 1983 until 2008. Although data coverage is rather high after 1980, the yearly country coverage is far from the maximum of 134 countries that report wages in the October Inquiry. Most countries reported between the middle of the 1980s and the turn of the millennium. Only five countries (Germany, Mauritius, Norway, Philippines, and Puerto Rico) report wages for all 26 years.

Theoretically, the approach of the ILO could result in an ideal database. Comparing wages for 161 occupations in 135 countries all over the world for a large period of time would promise an improvement in the analysis of wage growth and wage inequality. However, the October Inquiry database is far away from being ready to use for research purposes. The results of the survey are published without any correction or adjustment. Cleaning and correcting the data is a very time-consuming process. Moreover, as the reported wages differ, for instance, in reference time and in gender, wages are not comparable. To give few examples: Germany reports hourly, daily, or monthly minimum wages as an average for both sexes. China reports average yearly or monthly wages for men, women and averages for both sexes. Canada reports hourly minimum, maximum or averaged wages for men, women, and/or

[^1]both sexes. Table A. 1 shows the different reference time periods and the respective number of observations. Since the data is at this time neither comparable across countries nor within countries, or occupations, „the survey is one of the least widely used sources of cross country data in the world" (Freeman and Oostendorp 2000).

Since we are convinced that supplying comparable wage data for such a large number of countries yields an improvement for economic research, we transformed the October Inquiry into a usable and comparable form, which allows analyzing wage growth, wage gaps, and inequality in a comprehensive way. Yet, that required a comprehensive data correction and standardization procedure.

## 3. Data Corrections and Standardization Process

Since countries report the data in numerous ways, the October Inquiry wage observations are mostly not comparable. Neither within nor across countries, wages are reported consistently. Even within countries or for a particular occupation, wages are not comparable. Therefore, a considerably correction and standardization process is necessary.

### 3.1. Data structure

The data is very unbalanced and the reported wages differ in various dimensions. First, wages differ in the time they refer to. Within six different reported reference time periods (hourly, daily, fortnightly, weekly, monthly, and yearly wages), there are several other structures: for example minimum, average, and median wages. Germany, for instance, reports mostly monthly minimum wages from collective bargaining agreements, the United States report median wages for hours or years, the Netherlands maximum yearly wages, and India maximum daily wages. Altogether the database reports 33 different time periods. The reported time is in some cases specific to a particular country-occupation combination. Germany, for example, reports daily wages for only three occupations (miner in coalmining industry, miner in other underground industry, underground helper) and thus for only $1.7 \%$ of all German wage data. For most other occupations, monthly wages are reported.

On average, every country reports wages in four different time periods, in maximum in 16 periods, and in minimum one time period.

Second, there is no regularity in reporting wages with respect to gender. There are three gender categories: men and women (averaged wages for both sexes), men, and women. Yet, if two countries report the minimum monthly wage of a cook in the year of 2003, these two wages are hard to compare if they differ in the reported gender, since the gender wage gap poses a systematic bias in the comparison. The raw data contains 134 countries of which 98 report wages in all three gender categories, 21 report in two, and 15 in only one gender category.

Third, countries do not report the data continuously from 1983 until 2008. Moreover, even if countries reports wages for every year, wages were not necessarily reported for all 161 occupations. In fact, the database contains two types of gaps: time gaps and „occupation" gaps. Table 1 gives a first impression of the unbalanced structure of the raw data.

Table 1
Unbalanced Data: Number of Observations and Combinations

|  | Minimum | Average | Maximum |
| :---: | :---: | :---: | :---: |
| Observations |  |  |  |
| Country | 3 (France) | 1,641 | $\begin{aligned} & 13,481 \\ & \text { (Korea) } \end{aligned}$ |
| Year | $\begin{gathered} 4,100 \\ (1983) \end{gathered}$ | 8,651 | $\begin{aligned} & 13,024 \\ & (2006) \end{aligned}$ |
| Occupation | $460$ <br> (Railway steam-engine fireman) | 1,397 | $2,492$ <br> (Cook) |
| Time Period | $\begin{gathered} 2 \\ (\text { Per Week (Minimum)) } \end{gathered}$ | 6,816 | $\begin{gathered} 62,766 \\ \text { (Per Month (Average)) } \end{gathered}$ |
| Combinations |  |  |  |
| Year - Country | $\begin{gathered} 23 \\ (2008) \end{gathered}$ | 53 | 66 <br> (1987/1990) |
| Country - Year | 1 <br> (Croatia, Djibouti, etc.) | 10 | $26$ <br> (Germany, Norway, etc.) |
| Occupation - Country | $42$ <br> (Coalmining engineer) | 109 | $122$ <br> (Cook/Construction Carpenter) |
| Country - Occupation | 1 <br> (France) | 111 | $161$ <br> (United Kingdom, Romania) |
| Time Period - Country | $\begin{gathered} 1 \\ \text { (Several times) } \end{gathered}$ | 20 | $96$ <br> (Per Month, average) |
| Country - Time Period | $\begin{gathered} 1 \\ \text { (Several Countries) } \end{gathered}$ | 4 | 16 (Guyana, Saint Lucia) |

The data is very unbalanced. On average, every country reports 1,641 wages for 109 occupations in ten years. Every occupation is reported 1,397 times on average. There are 13,024 observations per year, and 6,816 wages reported per reference time. But, as Table 1 shows, the variations are large. Most countries reported wages in the years 1987 and 1990, and least countries report in 2008.

There are 96 countries which report wages in the most often used time period (Per Month, average). On average, 20 countries report wages for each of the 33 different time periods. These differences in reporting the data makes the comparison of the wages as they are released in the October Inquiry impossible.

### 3.2. Data corrections

The described differences in reporting the wages complicate the comparison of the data a lot. Yet, the differences affect the data in a systematic way so that standardization can be achieved. More problematic are miss-codifications and single events such as currency reforms for which the data must be corrected. Because the October Inquiry is published without any correction or adjustment, we perform an extensive cleaning procedure. Cleaning and correcting the data has been a very time-consuming process.

We identify unnatural growth in wages over time. For every country-occupation combination, we check wages in local currency that changed from one period to the following in an unnatural way and stayed on that level (what might result from a currency reform or a change in the reported time period) or return again on the former level in the next period (what could be result from an outlier, error or misscoded data). We find large irregularities in the data. In some cases, the hourly wage is as high as a monthly wage in the same country and year, or a wage that is labeled as a monthly wage is ten times higher than in comparable occupations. That makes it necessary to analyze detailed wage growth for every country-occupation combination for the whole period of time, using information from the footnotes the ILO gives to almost every single wage observation. We find a high need for adjustment and correct such irregularities by relabeling and redefining payment periods, or adjusting for currency reforms. In some cases no correction is possible, thus we have to drop the observations or the country as a whole. The countries we had to drop are marked with dots in Table A.2.

The countries from the European Monetary Union (EMU) changed their currency (from national currencies to the Euro) in 1999 or later. That makes comparison over time rather cumbersome. We therefore decide to convert the national currency into Euros for all observation of EMU countries before 1999 using the Euro conversion rate. Thus, the standardize wage is a Euro wage even if it refers to a year before the introduction of the Euro. We proceeded in the same way for all countries with currency reforms in the time period 1983-2008. Therefore, the standardized wages are in the current local currency of every country. Additionally, we keep the reported wage.

Although after the correction process, the reported wages are labeled in correctly, the data is far from being comparable within and across countries, or occupations. The wage data has to be transformed into a usable „standard form" in order to create a wage structure, that is based on comparable wages.

### 3.3. Standardization process

After the correction and cleaning procedure, the database contains wage data for 26 years, 112 countries, and 161 occupations. Wages are listed in more than thirty different time dimensions, for men, for women, or averages for both sexes. Thus, the task is to normalize the data in order to create one single comparable standard wage. We follow Freeman and Oostendorp (2000) and choose the average monthly wage of a man as a standard form, which is the most common form of the reported wage. The standardization procedure assumes that all deviations from an observed average systematic effect are random for the all observations.

We start the standardization by simplifying the reported time periods. We are aware of the risk of losing information, for example, if some occupations are systematically paid for a particular time period. But there is not enough variation in the data to keep the more than thirty time periods. Thus, we multiply a weekly wage with factor 4.33, a fortnightly wage with factor 2.16 and divide a yearly wage through 12 to transform the data into monthly wages.

The standardization process requires that the reported time periods are not specific to a particular country. If only a few countries report their wages in a particular time period (for example prevailing hourly wages) or one country dominates a particu-
lar time period (only the United States report median hourly wages), the effect of these time period can not be estimated independently from the country effect. We therefore adjust and merge these reference time periods with a closely related time period. As there are, for example, only few observations that are labeled "Per hour worked. Minimum." the observations are grouped to "Per hour. Minimum.". That reduces the number of subcategories of time periods for which wages are reported to 18 . Table A. 1 shows the number of observations reported per time period for the raw and the cleaned data.

As mentioned above, we choose the most common form of the reported wage as the standard form: the average monthly wage of a man. Although the average monthly wage of a man is the most common form, it applies only to ten percent of the data. We nevertheless dare to undertake the standardization procedure that translates the wage of each country-year-occupation observation, which is reported for another time period and/or gender, to man's average monthly wages. Thus, controlling for country, year, and occupation effects allows us to compute factors, that contain the deviation from any time period-gender combination to average monthly wages of men.

Suppose each wage observation $W$ (in logs) is the sum of the (unobserved) log wage in standard form (monthly average (ma) for a man (men)), $W^{*}$, and an adjustment coefficient $W^{a}$. The adjustment coefficient contains the deviation of the observed log wage from the standard wage, $W^{*}$. The observed wage, $W$, can then be described as:

$$
\begin{equation*}
W_{j, t, o, t d, s}=W_{j, t, o, m a, \text { men }}^{*}+W_{t d, s}^{a}+v_{j, t, o, t d, s}, \tag{1}
\end{equation*}
$$

where $j$ refers to the country $(j=1, \ldots, 112), t$ is the year $(t=1983, \ldots, 2008)$, $o$ denotes the occupation $(o=1, \ldots, 161)$, $t d$ is the time period (for example Per Hour. Average., $(t d=1, \ldots, 18)$ ), $s$ denotes the sex ( $s=$ average,men,women), and $v_{j, t, o, t d, s}$ is an error term.

The vector of the adjustment coefficients, $W_{t d, s}^{a}$, contains the conversion factors of any given time period-gender structure to average monthly wages for man for any given country-year-occupation observation. The adjustment coefficients can be calculated if we know the differences of the reported wages for a particular time period and gender to the standard wages, thereby controlling for country, year, and
occupation effects. The difference between the time period and the gender for the reported wages and the standard wage can be derived from a regression framework that explains wages by the time period, gender, occupation, year, and country effects. We chose country-year pairs instead of average time effects over all countries and average country effects over all years.

The regression equation for the observed wage is given by equation (2) and estimated taking into account that the residuals are heteroscedastic (Wooldridge 2001). We cluster around country-occupation pairs.

$$
\begin{equation*}
W_{j t, t d, o, s}=D_{t d} \alpha_{t d}+D_{s} \alpha_{s}+D_{o} \alpha_{o}+D_{j t} \alpha_{j t}+v_{j t, t d, o, s}, \tag{2}
\end{equation*}
$$

where $D_{t d}$ is a row vector of eighteen time periods, with "Per month. Average." being the reference period. $D_{s}$ is a row vector of the gender dummies, where we chose "Men" as reference. $D_{o}$ denotes a row vector of 161 occupation dummies, taking the cook as reference, which is the occupation with the most observations. Finally, $D_{j t}$ contains 1184 country-year dummies. We chose the United States in 2006 as reference. The vectors $\alpha_{t d}, \alpha_{s}, \alpha_{o}, \alpha_{j t}$ give the systematic deviation of the observed wages from the standard wage, respectively.

## 4. Regression Results and Interpretation of the Coefficients

The results of standardization process allow to transform the October Inquiry into a form that makes cross-country comparisons possible. Moreover, interpreting the estimated coefficient makes it possible to analyze the differences in wages explained by the reference time periods and the gender wage gaps. In the this section we present and discuss our results.

Equation (2) explains a great part of the variation in the data. The adjusted $R^{2}$ is 0.99 . That confirms that the standardization procedure is not afflicted with large errors. The dummy variables have the correct sign and are of the right magnitude. Since we estimate the regression equation in logs, we use the exponential function to compute adjustment coefficients of the reference period and the gender effect. These adjustment coefficients are used to convert the observed wages in their standard form, as they contain the difference of the observed from the standard wage. If, for example, the minimum wage per hour applying to women in a particular coun-
try, year and occupation is converted to the standard monthly wage of a man, the observed wage must be multiplied by 188.757 for Per hour. Minimum. and 1.187 for women, which yields an adjustment coefficient of 224.055.

The regression estimates and the resulting adjustment coefficients of the different reported time periods are presented in Table 2. Column one gives the regression results and standard errors of equation (2), column two refers to the computed adjustment coefficients. The adjustment coefficient is one, if the time has the standard form Per Month. Average. If the observed wage is not of that standard form, it has to be multiplied with the adjustment coefficient to yield the average monthly wage. The coefficients suggest that our adjustments are plausible. We find, for example, an adjustment coefficient that is lower than one for maximum monthly wages and higher than one for minimum monthly wages.

Table 2
Time Periods Regression Results and Adjustment Coefficients

| Time | Regression Coefficient | Adjustment Coefficient |
| :---: | :---: | :---: |
| Per hour. | $\begin{gathered} -4.911 * * * \\ (0.020) \end{gathered}$ | 135.744 |
| Per hour. Adjusted. | $\begin{gathered} -5.039 * * * \\ (0.024) \end{gathered}$ | 154.327 |
| Per hour. Average. | $\begin{gathered} -5.183 * * * \\ (0.017) \end{gathered}$ | 178.242 |
| Per hour. Maximum. | $\begin{gathered} -5.054 * * * \\ (0.025) \end{gathered}$ | 156.571 |
| Per hour. Minimum. | $\begin{gathered} -5.240 * * * \\ (0.019) \end{gathered}$ | 188.757 |
| Per hour paid for. | $\begin{gathered} -5.130 * * * \\ (0.024) \end{gathered}$ | 169.098 |
| Per hour worked. | $\begin{gathered} -4.916 * * * \\ (0.019) \end{gathered}$ | 136.483 |
| Per day. | $\begin{gathered} -2.423 * * * \\ (0.050) \end{gathered}$ | 11.285 |
| Per day. Adjusted. | $\begin{gathered} -3.047 * * * \\ (0.051) \end{gathered}$ | 21.056 |
| Per day. Average. | $\begin{gathered} -3.211 * * * \\ (0.038) \end{gathered}$ | 24.802 |
| Per day. Maximum. | $\begin{gathered} -3.527 * * * \\ (0.044) \end{gathered}$ | 34.024 |
| Per day. Minimum. | $\begin{gathered} -4.008 * * * \\ (0.037) \end{gathered}$ | 55.029 |
| Per month. | $\begin{gathered} 0.169 * * * \\ (0.005) \end{gathered}$ | 0.844 |
| Per month. Adjusted. | $\begin{gathered} 0.060^{* * *} \\ (0.017) \end{gathered}$ | 0.942 |
| Per month. Average. | $\begin{gathered} 0.000 \\ (0.000) \end{gathered}$ | 1.000 |
| Per month. Maximum. | $\begin{gathered} 0.234 * * * \\ (0.030) \end{gathered}$ | 0.792 |
| Per month. Minimum. | $\begin{gathered} -0.116 * * * \\ (0.016) \end{gathered}$ | 1.123 |
| Per month. Prevailing. | $\begin{array}{r} -0.030 \\ (0.027) \end{array}$ | 0.971 |
| All time periods relative to "Per month. Average". <br> All estimations include country-year, sex- and occupation-dummies. |  |  |

In the regression analysis, the time adjustment coefficient is based on an averaged effect over all countries. We are aware of the fact, that people in less developed countries might, for instance, work more than 20 days a month. Since there is not enough variation in the data, it is not possible to estimate time coefficients depending on the development level of countries. ${ }^{2}$

The gender adjustment coefficients presented in Table 3 reveal that men's wages are about 16 percent higher than those of women and about three percent higher than those o the average for men and women. In our analysis, the gender coefficient is also constant over time and across different groups of countries. In future work, we will have a closer look at the changes of the gender factor over time and between different country group and for the whole sample. We expect that the gender-gap is not constant over time. Moreover, there are possibly pronounced differences between countries that we expect to vary with the level of development.

Table 3
Gender Coefficients

| Sex | Regression Coefficient | Adjustment Coefficient |
| :--- | :---: | :---: |
| Men and Women | $-0,030 * * *$ | 1,031 |
| Men | $(0,006)$ | 1,000 |
| Women | $-0,172 * * *$ | 1,187 |
|  | $(0,004)$ |  |
| All time periods relative to Men |  |  |
| All estimations include country-year and occupation dummies. |  |  |

Applying the appropriate adjustment factors to all observed wages yields standard wages for all country-year-occupation combinations. Many country-year-occupation combinations occur more than once in the data, because the countries report more than one reference period (e.g. Per Month. Average and Per Month. Minimum.) or because countries report the wage for more than one gender for a particular yearoccupation combination. We keep the standardized wage with the shortest way to average monthly wages for men, but take into account the precision of the estimated parameter. We drop the other country-year-occupation observation and end-up with a dataset that holds only one observation for each country-year-occupation combination. That reduces the number of observations in our dataset to 93,535 but leaves the number of countries and year-country-occupation combinations unchanged.

[^2]Table 4 shows the change in the data that results from the four steps of modification that we have conducted.

Table 4
Construction of the data sets

| Data Set | Observations | Country-Year-Occupation | Number of Countries |
| :--- | :---: | :---: | :---: |
| Raw Data | 224,570 | 109,651 | 134 |
| Cleaned Data | 182,786 | 93,535 | 112 |
| Standardized Data | 93,535 | 93,535 | 112 |
| Imputed Data | 147,016 | 147,016 | 112 |

## 5. A First Glance: Some Descriptive Statistics

To give a first impression of the data, we present some summary statistics in this section. The data is still very unbalanced, as it contains many gaps in time and for particular occupations. In combination with the different dimensions of the data, presenting descriptive statistics is not easy.

## Table 5

Descriptive Statistics

|  | Growth Rate |  |  | Wage Gap |  |  | Variation Coefficient |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Year | Mean | Min. | Max. | Mean | Min. | Max. | Mean | Min. | Max. |
| Whole sample, varying number of countries |  |  |  |  |  |  |  |  |  |
| 1986 | . | . | . | 15.54 | 1.05 | 59.46 | 0.66 | 0.04 | 1.48 |
| 1996 | 12.15 | -2.93 | 64.62 | 12.24 | 1.92 | 74.58 | 0.52 | 0.18 | 1.30 |
| 2006 | 6.84 | -1.85 | 34.54 | 14.29 | 1.16 | 66.34 | 0.58 | 0.05 | 1.13 |
| OECD countries, varying number of countries |  |  |  |  |  |  |  |  |  |
| 1986 | . | . | . | 8.97 | 1.70 | 15.07 | 0.48 | 0.14 | 1.07 |
| 1996 | 5.99 | 2.51 | 16.78 | 6.08 | 1.92 | 15.02 | 0.37 | 0.18 | 0.60 |
| 2006 | 3.65 | -1.85 | 23.54 | 7.08 | 2.11 | 13.26 | 0.45 | 0.22 | 0.70 |
| Non-OECD countries, varying number of countries |  |  |  |  |  |  |  |  |  |
| 1986 | . |  |  | 18.80 | 1.05 | 59.46 | 0.74 | 0.04 | 1.48 |
| 1996 | 19.13 | -2.93 | 64.62 | 17.30 | 2.71 | 74.58 | 0.64 | 0.25 | 1.30 |
| 2006 | 10.60 | -1.09 | 34.54 | 22.57 | 1.16 | 66.35 | 0.74 | 0.05 | 1.13 |

Growth rate: un-weighted average of wage-ratios in $t$ over in $t-1$ for all
occupations within a country (in Percent).
Wage gap: highest wage over lowest wage within a country.
Variation coefficients: Standard Deviation/Mean within each country.

We decided to give un-weighted averages of annual growth rates, wage gaps (highest wage over lowest wage), and within country wage variation coefficients for the whole sample and a split for OECD and Non-OECD countries for 1986, 1996, and 2006. The results are presented in Table 5.

The annual growth rate of nominal wages falls from $12.2 \%$ on average between 1986 and 1996 to $6.8 \%$ between 1996 and 2006. In parts, this fall reflects declining inflation. This can be seen by contrasting the results from Table 5 with the first column of Table 6, that shows the average growth rates of three occupations in US Dollar. ${ }^{3}$ These growth rates are far smaller, because denominating the wages in US Dollar controls partly for inflation in all countries except the United States. The differences in the average growth rates shown in Table 5 are also to some extend the result of changing composition of the sample over the years. Yet, the fall in the growth rate does not seem to be driven by outliers. The fall in the growth rats of nominal wages is apparent for both groups, OECD-countries and Non-OECD countries by a similar factor. The wage structure as a whole is robust against these changes in the sample. Neither the wage gap, i.e. the ratio of the largest over the smallest wage in each country, nor the variation coefficient varies much over time. Note that both measures are by construction not affected by inflation.

At this very aggregated level, wage income does not seem to have increased on average. The differences between different occupations within the countries have remained stable. The wage gap has increased slightly, whereas the variation coefficient has decreased. Moreover, splitting the countries between OECD and NonOECD countries reveals no different pattern for the two groups. The growth rates difference between the OECD and Non-OECD countries, however, seem to indicate decreasing differences between the countries.

In Table 6, we look at the differences between occupations. We therefore chose three occupations with coverage as representative for three different skill levels. The low skilled Waiter and the high skilled Physician show very different pattern on the average for all countries.

[^3]Table 6
Descriptive Statistics: Occupational Wage Gap

|  | Growth Rate |  |  | Wage Gap | Variation Coefficient |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Year | Mean | Min. | Max. | Mean | Mean |
| Occupation: Low skilled, Medium Skilled, High Skilled |  |  |  |  |  |
| Low skilled (Waiter) |  |  |  |  |  |
| 1986 | . | . | . | 77.49 | 1.09 |
| 1996 | 4.79 | -5.47 | 11.90 | 48.89 | 1.00 |
| 2006 | 3.06 | -0.81 | 10.23 | 39.14 | 1.00 |
| Medium Skilled (Cook) |  |  |  |  |  |
| 1986 | . | . | . | 95.49 | 0.97 |
| 1996 | 5.93 | -5.65 | 13.53 | 45.59 | 1.02 |
| 2006 | 3.01 | -8.77 | 9.77 | 40.94 | 0.97 |
| High Skilled (Gerneral Physician) |  |  |  |  |  |
| 1986 | . | . | . | 49.17 | 0.85 |
| 1996 | 3.34 | -9.66 | 11.45 | 72.24 | 1.09 |
| 2006 | 4.27 | -7.76 | 15.90 | 135.20 | 0.97 |

Growth rate: un-weighted average of wage-ratios in $t$ over in $t-1$ within each occupation.
Wage gap: highest wage over lowest wage within each occupation.
Variation coefficients: Standard Deviation/Mean within each occupation.
While the growth rate and the wage gap decrease for the Waiter, both increase for the high skilled Physician. The variation coefficients also reveals that the differences between the countries decrease for the low skilled Waiter, but increase for the high skilled Physician.

To analyze differences in wages between the countries in a more comprehensive way, we regress occupation-dummies on the log standardized using three regression frameworks. The first includes all countries contained in the dataset ${ }^{4}$, the second only OECD-countries, and the third Non-OECD-Countries. ${ }^{5}$ We control for country-year-effects using the United States in the year 2006 as benchmark, as we did in equation (2). We compare the results of the three regression frameworks with those of (2). The results are presented in Tables A. 3 and A.4.

Every coefficient has to be interpreted in relation to the benchmark occupation: the Cook. The occupation coefficient is one for Cook, it is larger than one if the average wage of an occupation is lower than the wage of a Cook, and it is lower than one

[^4]if the average wage of an occupation is higher than the wage of a Cook. We find, for instance, that the wage of a waiter is on average 19 percent lower than the wage of a Cook. But, in OECD-countries it is nine percent lower, and 23 percent lower in Non-OECD-countries. The wage of a salesperson in wholesale is on average seven percent higher than the wage of a Cook, 17 percent in OECD- and about four percent in Non-OECD-countries. Looking at a salesperson in retail trade, we find, that the wage is lower than the wage of a Cook. On average it is 12 percent lower, three percent in OECD- and 16 percent Non-OECD countries. The best-paid occupation is the airtransport pilot, whose wage is almost five times higher than that of a Cook. The worst paid occupation is a laborer in the spinning and weaving industry, with an average wage which almost 30 percent lower, nine in OECD- and more than 50 percent in Non-OECD-Countries.

## 6. Imputation

The standardization step leaves us with a dataset that contains comparable data within and between countries. Yet, the data is unbalanced. There are gaps with respect to occupations, i.e. not all countries report wages of all occupation for every year, and with respect to time, i.e. not all countries report wages for every year. Often, countries report every other year, but for some countries no pattern can be found. We decided to fill some gaps in order to base the cross-section comparisons on a larger sample. To make sure that we do not impose a structure on the data, we fill only small gaps and only use the within-country variation for the imputations.

With respect to the occupations gaps, we impute only those missing wages, for which we can compute the coefficients of the occupation dummies with sufficient precision. The coefficients result from a multivariate regression similar to the one employed in the standardization process. For imputation we regressed the standardized wage on occupation and time dummies for each country separately. We assume that the wage structure does not change much over time and we impute the missing wages by using the occupation dummy variables. They reflect the wage pattern averaged over all years. Since the cook is the most reported occupation, we choose it as benchmark and compute the coefficients of the occupation dummies by using the exponential function. In some cases the wage of the cook is not reported for every year, we therefore we choose the most often reported occupation as a benchmark
instead. With respect to time, we decide to fill only one-year gaps. Thus, we know the wage in the year before and in the year after the missing. We use linear projections to impute the missing wages. That increases the number of observations from 93,535 to 147,016 . Our imputation procedure does not change the structure of the data, neither their time structure nor the wage pattern with respect to occupations. The occupation coefficients using the imputed data are shown in columns 2, 3, and 4 in Tables A. 3 and A.4. They have the same structure as the coefficients using the standardized data (column 1). Thus, the imputation process of the data does not change the structure of the reported wages.

The imputed dataset contains standardized wages for up to 161 occupations from 49 industries in 112 countries between 1983 and 2008. The data is ready to use for many applications and relatively easy to adjust for others. The standard wage is given in local currency and in US-Dollar.

## 7. Conclusion

The October Inquiry is a rarely used database. The correction process applied in this paper is extensive and time-consuming. Almost every country had to be corrected separately. We corrected the data for typos, outliers, and mislabeled observations, and reviewed the data for currency reforms or unnatural high inflation. Since wages were reported in numerous different time periods, we had to reduce the number of reference times. The required standardization approach is complex, but does not change the structure of the data. Neither does the imputation.

Our approach yields to a comprehensive database, that allows analyzing worldwide comparable wage data for a large number of countries and occupations. Moreover, the estimated adjustment factors can be interpreted as occupational and gender wage gaps. We find that wages differ by about 15 percent between men and women, and by three percent between men and the averaged wage of men and women. Our approach assumes that the wage gap is independent from the level of development of the different countries. In future work, we will analyze how restrictive this assumption is and whether it biases our results.

A first glance at the data points to decreasing differences between the countries and stable differences among the occupations within the countries. The falling differ-
ences between the countries seem to be driven by falling differences in the wages of low skilled occupations. The wages of the high skilled, in contrast, differ between the countries. However, an in-depth study of the evolution of the wage structure has to follow.

## References

Freeman, Richard B. and Remco Oostendorp, (2001). The Occupational Wages around the World data file, International Labour Review, Vol. 140, No.4.
Freeman, Richard B. and Remco Oostendorp, (2000). Wages Around the World: Pay Across Occupations and Countries, NBER Working Papers No. 8058, National Bureau of Economic Research, Inc.
Wooldridge, Jeffrey (2001). The Econometric Analysis of Cross Section and Panel Data. Cambridge: The MIT Press.

## A. Appendix

Table A. 1
Time Periods: Raw and Cleaned Data

| Time Period | Observations |  | Time Period | Observations |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | Raw Data | Cleaned Data |  | Raw Data | Cleaned Data |
| Per hour |  |  | Per fortnight |  |  |
| Per hour. | 7,701 | 6,069 | Per fortnight. | 320 |  |
| Per hour. Adjusted. |  | 4,444 | Per fortnight. Average. | 195 |  |
| Per hour. Average. | 9,568 | 6,708 | Per fortnight. Maximum. | 14 |  |
| Per hour. Maximum. | 1,279 | 795 | Per fortnight. Minimum. | 303 |  |
| Per hour. Median. | 1,513 |  | Per fortnight. Prevailing. | 9 |  |
| Per hour. Minimum-maximum. | 204 |  | Per month |  |  |
| Per hour. Minimum. | 6,992 | 5,412 | Per month. | 58,263 | 49,406 |
| Per hour. Prevailing. | 747 |  | Per month. Adjusted. |  | 25,827 |
| Per hour paid for. | 7,209 | 3,111 | Per month. Average. | 64,338 | 54,249 |
| Per hour paid for. Hours paid for. | 16 |  | Per month. Maximum. | 2,948 | 1,085 |
| Per hour paid for. Maximum. | 10 |  | Per month. Median. | 1,460 |  |
| Per hour paid for. Minimum. | 10 |  | Per month. Minimum-maximum. | 107 |  |
| Per hour worked. | 3,251 | 3,092 | Per month. Minimum. | 19,216 | 13,038 |
| Per hour worked. Maximum. | 38 |  | Per month. Prevailing. | 2,255 | 1,440 |
| Per hour worked. Minimum. | 38 |  | Per year |  |  |
| Per day |  |  | Per year. | 371 |  |
| Per day. | 1,116 | 804 | Per year. Average. | 1,199 |  |
| Per day. Adjusted. |  | 426 | Per year. Maximum. | 674 |  |
| Per day. Average. | 2,205 | 1,707 | Per year. Median. | 824 |  |
| Per day. Maximum. | 1,790 | 812 | Per year. Minimum. | 794 |  |
| Per day. Median. | 32 |  | Per year. Prevailing. | 169 |  |
| Per day. Minimum. | 6,193 | 4,302 |  |  |  |
| Per day. Prevailing. | 113 |  |  |  |  |
| Per week |  |  |  |  |  |
| Per week. | 12,492 |  |  |  |  |
| Per week. Average. | 3,15 |  |  |  |  |
| Per week. Maximum. | 172 |  |  |  |  |
| Per week. Median. | 2,573 |  |  |  |  |
| Per week. Minimum-maximum. | 13 |  |  |  |  |
| Per week. Minimum. | 2,442 |  |  |  |  |
| Per week. Prevailing. | 141 |  |  |  |  |
| Cleaned Data: yearly wages divided through 12, weekly wages multiplied with factor 4.33, and fortnightly wages with factor 2.16. |  |  |  |  |  |

Table A. 2
Observations per Country: Cleaned Data, Standardized Data, Imputed Data

| Country | Cleaned Data | Standardized Data | Imputed Data | Country | Cleaned Data | Standardized Data | Imputed Data | Country | Cleaned Data | Standardized Data | Imputed Data |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Algeria | 1,110 | 1,019 | 2,812 | Ethiopia | 47 | 47 | 58 | Nepal | 114 | 108 | 170 |
| Angola | . |  | . | Falkland Islands (Malvinas) | . | . | . | Netherlands | 700 | 386 | 408 |
| Antigua and Barbuda |  |  |  | Fiji |  |  |  | Netherlands Antilles | 533 | 301 | 576 |
| Argentina | 474 | 474 | 1,150 | Finland | 7,342 | 2,670 | 3,384 | New Caledonia | 65 | 65 | 65 |
| Australia | 3,466 | 1,825 | 2,718 | France |  |  |  | Nicaragua | 565 | 565 | 1,000 |
| Austria | 5,250 | 2,655 | 3,000 | French Polynesia | 326 | 326 | 360 | Nigeria |  |  |  |
| Azerbaijan | 827 | 470 | 676 | Gabon | 406 | 247 | 776 | Norway | 1,089 | 758 | 1,482 |
| Bahamas | 502 | 334 | 927 | Germany | 3,990 | 3,990 | 4,134 | Pakistan | 1,379 | 773 | 1,106 |
| Bahrain | 1,462 | 991 | 2,622 | Gibraltar | 429 | 246 | 468 | Papua New Guinea | 480 | 418 | 882 |
| Bangladesh | 1,985 | 1,201 | 1,960 | Grenada | 406 | 343 | 420 | Peru | 2,442 | 1,042 | 2,160 |
| Barbados | 1,223 | 675 | 949 | Guam | 200 | 110 | 450 | Philippines | 960 | 842 | 2,520 |
| Belarus | 2,199 | 658 | 715 | Guatemala | 407 | 217 | 596 | Poland | 2,450 | 1,057 | 1,771 |
| Belgium | 874 | 848 | 1,188 | Guyana | 933 | 869 | 2,227 | Portugal | 7,580 | 1,590 | 3,312 |
| Belize | 1,913 | 1,135 | 1,365 | Honduras | 2,546 | 1,425 | 1,950 | Puerto Rico | 216 | 192 | 1,104 |
| Benin | 766 | 413 | 1,125 | Hong Kong | 1,639 | 830 | 1,534 | Romania | 8,365 | 3,253 | 3,381 |
| Bermuda | 410 | 359 | 1,059 | Hungary | 3,432 | 1,782 | 2,086 | Russian Federation | 1,482 | 623 | 1,342 |
| Bolivia | 2,328 | 1,254 | 1,898 | Iceland | 809 | 325 | 1,274 | Rwanda | 845 | 845 | 1,008 |
| Botswana | 152 | 152 | 184 | India | 2,584 | 1,324 | 1,761 | Saint Kitts and Nevis | . | . | . |
| Brazil | 957 | 383 | 1,206 | Indonesia | 832 | 415 | 1,302 | Saint Lucia | . | . |  |
| Brunei Darussalam |  |  |  | Ireland | 30 | 30 | 30 | Saint Pierre and Miquelon | . | . | . |
| Bulgaria | 508 | 233 | 122 | Isle of Man |  |  |  | Saint Vincent and the Grenadines |  |  |  |
| Burkina Faso | 775 | 677 | 1,276 | Italy | 3,479 | 3,479 | 3,672 | San Marino | 293 | 293 | 404 |
| Burundi | 585 | 567 | 810 | Japan | 1,384 | 939 | 1,248 | Senegal | 73 | 73 | 73 |
| Cameroon | 524 | 524 | 1,190 | Jordan | 3,106 | 1,868 | 2,907 | Serbia and Montenegro | 159 | 159 | 159 |
| Canada | 2,365 | 1,249 | 1,860 | Kazakhstan | 838 | 347 | 351 | Seychelles | . | . | . |
| Cape Verde | 161 | 159 | 160 | Kenya | 254 | 157 | 176 | Sierra Leone |  |  |  |
| Central African Republic | 723 | 723 | 1,276 | Korea | 13,445 | 3,247 | 3,792 | Singapore | 3,838 | 2,060 | 3,473 |
| Chad | 913 | 732 | 1,122 | Kuwait | 372 | 128 | 128 | Slovakia | 4,336 | 1,670 | 2,041 |
| Chile | 579 | 441 | 720 | Kyrgyzstan | 550 | 216 | 396 | Slovenia | 366 | 303 | 728 |
| China | 1,593 | 976 | 1,834 | Latvia | 3,372 | 1,195 | 1,480 | Sudan | . | . | . |
| Colombia | 223 | 223 | 417 | Lesotho | 204 | 194 | 230 | Suriname |  |  |  |
| Comoros | 987 | 786 | 1,404 | Liberia | 59 | 54 | 86 | Sweden | 2,247 | 1,192 | 1,898 |
| Costa Rica | 3,185 | 1,661 | 2,415 | Lithuania | 1,136 | 363 | 705 | Thailand | 3,521 | 1,008 | 1,400 |
| Croatia | 119 | 119 | 119 | Luxembourg | 456 | 165 | 267 | Togo | 216 | 213 | 336 |
| Cuba | 2,543 | 1,199 | 1,460 | Madagascar | 1,221 | 841 | 1,264 | Trinidad and Tobago | 758 | 680 | 1,304 |
| Cyprus | 7,798 | 2,348 | 2,852 | Malawi | 934 | 676 | 1,350 | Tunisia |  |  |  |
| Czech Republic | 2,931 | 1,768 | 2,208 | Malaysia | 583 | 251 | 1,106 | Turkey | 277 | 153 | 330 |
| Czechoslovakia | 875 | 834 | 1,120 | Maldives | 66 | 36 | 36 | Ukraine | 152 | 152 | 300 |
| Côte d'Ivoire | 992 | 787 | 1,738 | Mali |  |  |  | United Kingdom | 7,371 | 1,914 | 3,864 |
| Denmark | 1,804 | 1,010 | 1,770 | Mauritius | 3,649 | 1,987 | 2,964 | United States | 3,468 | 2,501 | 3,850 |
| Djibouti | 48 | 48 | 48 | Mexico | 4,008 | 1,707 | 2,717 | Uruguay | 853 | 489 | 572 |
| Egypt | 916 | 678 | 1,624 | Mongolia | 44 | 44 | 44 | Venezuela | 1,133 | 975 | 1,540 |
| El Salvador | . |  |  | Mozambique | 350 | 350 | 444 | Virgin Islands (US) | . | . | . |
| Eritrea | 366 | 279 | 375 | Myanmar |  |  |  | Zambia | . | . | . |
| Estonia | 1,119 | 525 | 705 | Namibia | 66 | 50 | 105 | Total | 182,786 | 93,535 | 147,016 |

Table A. 3
Coefficients for Low Skilled and High Skilled Occupations

| Occupation | Average (1) | Average (2) | OECD | Non-OECD | Occupation | Average (1) | Average (2) | OECD | Non-OECD |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Low Skilled |  |  |  |  | High Skilled |  |  |  |  |
| Deep-sea fisherman | 1.023 | 1.018 | 0.955 | 1.045 | Accountant | 0.382 | 0.385 | 0.471 | 0.352 |
| Dockworker | 0.867 | 0.905 | 0.777 | 0.974 | Air traffic controller | 0.397 | 0.419 | 0.385 | 0.431 |
| Field crop farm worker | 1.322 | 1.385 | 1.131 | 1.549 | Air transport pilot | 0.229 | 0.209 | 0.258 | 0.187 |
| Forestry worker | 1.157 | 1.207 | 1.071 | 1.277 | Aircraft engine mechanic | 0.516 | 0.510 | 0.602 | 0.463 |
| Inshore (coastal) maritime fisherman | 1.054 | 1.028 | 0.966 | 1.056 | Auxiliary nurse | 0.924 | 0.949 | 0.849 | 1.006 |
| Labourer (construction) | 1.171 | 1.239 | 0.973 | 1.394 | Book-keeping machine operator | 0.628 | 0.683 | 0.717 | 0.678 |
| Labourer (printing) | 1.145 | 1.202 | 0.963 | 1.332 | Chemical engineer | 0.440 | 0.403 | 0.427 | 0.384 |
| Labourer (manufacturing) | 1.108 | 1.159 | 0.964 | 1.265 | Chemistry technician | 0.662 | 0.656 | 0.619 | 0.668 |
| Labourer (manufacturing. chemical) | 1.189 | 1.225 | 0.999 | 1.357 | Coalmining engineer | 0.485 | 0.470 | 0.495 | 0.431 |
| Labourer (Iron and steel) | 1.039 | 1.102 | 0.924 | 1.215 | Computer programmer (insurance) | 0.453 | 0.463 | 0.495 | 0.442 |
| Labourer (manufacture of machinery) | 1.145 | 1.220 | 1.031 | 1.341 | Computer programmer (public administration) | 0.529 | 0.562 | 0.520 | 0.580 |
| Labourer (Spinning. weaving. finishing textile) | 1.307 | 1.390 | 1.090 | 1.573 | Dentist (general) | 0.469 | 0.472 | 0.421 | 0.484 |
| Labourer (electric light and power) | 1.038 | 1.116 | 0.933 | 1.206 | Electronics engineering technician | 0.660 | 0.686 | 0.609 | 0.727 |
| Logger | 1.196 | 1.233 | 1.019 | 1.375 | Fire-fighter | 0.851 | 0.890 | 0.737 | 0.973 |
| Office clerk | 0.888 | 0.943 | 0.821 | 1.005 | First-level education teacher | 0.664 | 0.718 | 0.617 | 0.768 |
| Packer (Slaughtering) | 1.147 | 1.164 | 1.006 | 1.254 | Flight operations officer | 0.420 | 0.406 | 0.399 | 0.400 |
| Packer (manufacturing. chemical) | 1.084 | 1.182 | 0.914 | 1.346 | Forest supervisor | 0.821 | 0.830 | 0.758 | 0.863 |
| Plantation worker | 1.348 | 1.389 | 1.051 | 1.575 | General physician | 0.378 | 0.381 | 0.352 | 0.390 |
| Postman | 0.953 | 1.003 | 0.847 | 1.073 | Government executive official: a) | 0.472 | 0.453 | 0.483 | 0.441 |
| Railway vehicle loader | 1.052 | 1.110 | 0.961 | 1.175 | Government executive official: b) | 0.435 | 0.446 | 0.476 | 0.438 |
| Room attendant or chambermaid | 1.216 | 1.305 | 1.177 | 1.369 | Government executive official: c) | 0.527 | 0.582 | 0.527 | 0.604 |
| Telephone switchboard operator | 0.811 | 0.871 | 0.820 | 0.895 | Journalist | 0.526 | 0.536 | 0.522 | 0.541 |
| Ticket seller (cash desk cashier) | 0.847 | 0.891 | 0.788 | 0.946 | Kindergarten teacher | 0.845 | 0.858 | 0.736 | 0.925 |
| Tree feller and bucker | 1.222 | 1.232 | 0.991 | 1.373 | Mathematics teacher (second level) | 0.560 | 0.594 | 0.547 | 0.618 |
| Underground helper. loader | 0.849 | 0.868 | 0.754 | 0.939 | Mathematics teacher (third level) | 0.444 | 0.459 | 0.445 | 0.466 |
| Waiter | 1.174 | 1.230 | 1.094 | 1.300 | Petroleum and natural gas engineer | 0.349 | 0.328 | 0.400 | 0.286 |
|  |  |  |  |  | Power distribution and transmission e | 0.389 | 0.382 | 0.447 | 0.358 |
|  |  |  |  |  | Power-generating machinery operator | 0.655 | 0.674 | 0.653 | 0.675 |
|  |  |  |  |  | Refuse collector | 1.150 | 1.179 | 0.997 | 1.285 |
|  |  |  |  |  | Ship's chief engineer | 0.489 | 0.453 | 0.468 | 0.434 |
|  |  |  |  |  | Supervisor or general foreman | 0.621 | 0.636 | 0.563 | 0.668 |
|  |  |  |  |  | Teacher (language. literature) (second level) | 0.560 | 0.589 | 0.544 | 0.612 |
|  |  |  |  |  | Teacher (language. literature) (third level) | 0.447 | 0.465 | 0.439 | 0.479 |

All estimations relative to occupation Cook; all estimations include country-year and occupation dummies.
,
Average (2), OECD, Non-OECD: Estimated on basis of imputed data; based on regression $W_{o, j t}=D_{o} \alpha_{o}+D_{j t} \alpha_{j t}+v_{o, j t}$

Table A. 4
Coefficients for Medium Skilled Occupations

| Occupation | Average (1) | Average (2) | OECD | Non-OECD | Occupation | Average (1) | Average (2) | OECD | Non-OECD |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Medium Skilled |  |  |  |  | Medium Skilled |  |  |  |  |
| Able seaman | 0.891 | 0.921 | 0.792 | 1.003 | Medical X-ray technician | 0.697 | 0.729 | 0.685 | 0.747 |
| Aircraft accident fire-fighter | 0.725 | 0.760 | 0.642 | 0.813 | Metal melter | 0.791 | 0.847 | 0.775 | 0.879 |
| Aircraft cabin attendant | 0.502 | 0.528 | 0.555 | 0.511 | Metalworking machine setter | 0.908 | 0.911 | 0.828 | 0.954 |
| Aircraft loader | 0.901 | 0.939 | 0.816 | 1.004 | Miner (Coalmining) | 0.710 | 0.724 | 0.662 | 0.754 |
| Airline ground receptionist | 0.633 | 0.643 | 0.718 | 0.613 | Miner (Other mining and quarrying) | 0.795 | 0.809 | 0.703 | 0.873 |
| Ambulance driver | 0.915 | 0.989 | 0.886 | 1.069 | Mixing- and blending-machine operator (manufactural chemical industry) | 0.836 | 0.848 | 0.744 | 0.907 |
| Automobile mechanic (Passenger transport) | 0.895 | 0.913 | 0.813 | 0.960 | Mixing- and blending-machine operator (other chemical industry) | 0.872 | 0.887 | 0.788 | 0.939 |
| Automobile mechanic (Repair of motor vehicles) | 0.924 | 0.925 | 0.869 | 0.948 | Motor bus driver | 0.879 | 0.933 | 0.828 | 0.989 |
| Baker (ovenman) | 1.075 | 1.111 | 0.928 | 1.223 | Occupational health nurse | 0.651 | 0.728 | 0.674 | 0.746 |
| Bank teller | 0.673 | 0.710 | 0.727 | 0.708 | Office clerk (Printing. publishing) | 0.831 | 0.886 | 0.753 | 0.961 |
| Bench moulder (metal) | 0.877 | 0.928 | 0.836 | 0.978 | Office clerk (Electric light and power) | 0.702 | 0.753 | 0.685 | 0.792 |
| Blast furnaceman (ore smelting) | 0.784 | 0.822 | 0.755 | 0.850 | Paper-making-machine operator | 0.892 | 0.926 | 0.774 | 1.026 |
| Book-keeper | 0.820 | 0.875 | 0.760 | 0.940 | Petroleum and natural gas extraction | 0.489 | 0.478 | 0.511 | 0.458 |
| Bookbinder (machine) | 0.967 | 0.983 | 0.821 | 1.073 | Physiotherapist | 0.708 | 0.723 | 0.754 | 0.694 |
| Bricklayer (construction) | 0.917 | 0.960 | 0.832 | 1.021 | Plantation supervisor | 0.906 | 0.942 | 0.862 | 0.981 |
| Building electrician | 0.842 | 0.844 | 0.746 | 0.891 | Plasterer | 0.909 | 0.972 | 0.827 | 1.045 |
| Building painter | 0.928 | 0.969 | 0.833 | 1.043 | Plumber | 0.879 | 0.913 | 0.775 | 0.983 |
| Bus conductor | 1.014 | 1.067 | 0.854 | 1.187 | Plywood press operator | 1.046 | 1.075 | 0.888 | 1.189 |
| Butcher | 1.019 | 1.036 | 0.904 | 1.109 | Post office counter clerk | 0.795 | 0.863 | 0.797 | 0.884 |
| Cabinetmaker | 1.047 | 1.095 | 0.957 | 1.166 | Printing pressman | 0.834 | 0.848 | 0.741 | 0.905 |
| Card- and tape-punching- machine operator (insurance) | 0.775 | 0.799 | 0.775 | 0.809 | Professional nurse (general) | 0.668 | 0.718 | 0.669 | 0.736 |
| Card- and tape-punching- machine operator (public administration) | 0.874 | 0.907 | 0.803 | 0.960 | Quarryman | 0.928 | 0.970 | 0.808 | 1.082 |
| Cash desk cashier | 1.038 | 1.058 | 1.020 | 1.083 | Railway engine-driver | 0.665 | 0.699 | 0.639 | 0.722 |
| Cement finisher | 0.927 | 0.955 | 0.805 | 1.042 | Railway passenger train guard | 0.769 | 0.792 | 0.769 | 0.765 |
| Clerk of works | 0.612 | 0.643 | 0.646 | 0.635 | Railway services supervisor | 0.629 | 0.615 | 0.605 | 0.587 |
| Clicker cutter (machine) | 1.170 | 1.224 | 1.070 | 1.301 | Railway signalman | 0.833 | 0.888 | 0.742 | 0.979 |
| Cloth weaver (machine) | 1.126 | 1.139 | 1.006 | 1.209 | Railway steam-engine fireman | 0.822 | 0.845 | 0.723 | 0.891 |
| Construction carpenter | 0.912 | 0.920 | 0.808 | 0.978 | Reinforced concreter | 0.925 | 0.950 | 0.813 | 1.020 |
| Constructional steel erector | 0.868 | 0.908 | 0.793 | 0.964 | Road transport services supervisor | 0.717 | 0.703 | 0.656 | 0.723 |
| Controlman | 0.595 | 0.575 | 0.585 | 0.546 | Salesperson (Wholesale trade) | 0.930 | 0.936 | 0.856 | 0.965 |
| Cook | 1.000 | 1.000 | 1.000 | 1.000 | Salesperson (Retail trade) | 1.125 | 1.125 | 1.021 | 1.178 |
| Dairy product processor | 0.977 | 1.005 | 0.881 | 1.072 | Sawmill sawyer | 1.070 | 1.120 | 0.931 | 1.234 |
| Derrickman | 0.691 | 0.722 | 0.674 | 0.739 | Sewing-machine operator | 1.229 | 1.250 | 1.109 | 1.329 |
| Electric power lineman | 0.699 | 0.732 | 0.679 | 0.752 | Ship plater | 0.862 | 0.898 | 0.818 | 0.932 |
| Electronic equipment assembler | 0.999 | 1.054 | 0.917 | 1.138 | Ship's steward (passenger) | 0.866 | 0.883 | 0.861 | 0.894 |
| Electronics draughtsman | 0.733 | 0.747 | 0.675 | 0.784 | Shoe sewer (machine) | 1.178 | 1.233 | 1.093 | 1.300 |
| Electronics fitter | 0.882 | 0.999 | 0.786 | 1.138 | Stenographer-typist (banks) | 0.639 | 0.664 | 0.743 | 0.635 |
| Farm supervisor | 0.863 | 0.862 | 0.869 | 0.839 | Stenographer-typist (insurance) | 0.740 | 0.751 | 0.792 | 0.734 |
| Furniture upholsterer | 1.048 | 1.076 | 0.965 | 1.129 | Stenographer-typist (public administration) | 0.844 | 0.869 | 0.860 | 0.879 |
| Garment cutter | 1.093 | 1.097 | 0.989 | 1.148 | Stenographer-typist (printing. publishing) | 0.807 | 0.834 | 0.839 | 0.838 |
| Grain miller | 0.938 | 0.956 | 0.887 | 0.991 | Stenographer-typist (wholesale trade) | 0.850 | 0.886 | 0.893 | 0.875 |
| Hand compositor | 0.880 | 0.901 | 0.756 | 0.988 | Stock records clerk | 0.852 | 0.899 | 0.867 | 0.913 |
| Hot-roller (steel) | 0.802 | 0.835 | 0.767 | 0.862 | Supervisor or general foreman | 0.525 | 0.497 | 0.529 | 0.468 |
| Hotel receptionist | 0.972 | 1.043 | 1.004 | 1.065 | Tanner | 1.072 | 1.105 | 0.947 | 1.184 |
| Insurance agent | 0.585 | 0.615 | 0.602 | 0.617 | Technical education teacher (second level) | 0.584 | 0.637 | 0.555 | 0.678 |
| Laster | 1.169 | 1.241 | 1.112 | 1.300 | Thread and yarn spinner | 1.135 | 1.162 | 1.022 | 1.242 |
| Leather goods maker | 1.234 | 1.267 | 1.084 | 1.257 | Urban motor truck driver | 0.952 | 0.992 | 0.958 | 0.991 |
| Long-distance motor truck driver | 0.876 | 0.885 | 0.859 | 0.882 | Veneer cutter | 1.066 | 1.102 | 0.915 | 1.213 |
| Loom fixer. tuner | 1.035 | 1.091 | 0.932 | 1.174 | Welder | 0.905 | 0.928 | 0.832 | 0.978 |
| Machine compositor | 0.835 | 0.840 | 0.737 | 0.896 | Wood grinder | 0.899 | 0.912 | 0.745 | 1.041 |
| Machinery fitter-assembler | 0.891 | 0.899 | 0.834 | 0.923 | Wooden furniture finisher | 1.092 | 1.125 | 0.959 | 1.213 |

All estimations relative to occupation Cook; all estimations include country-year and occupation dummies.
Average (2), OECD, Non-OECD: Estimated on basis of imputed data; based on regression $W_{o, j t}=D_{o} \alpha_{o}+D_{j t} \alpha_{j t}+v_{o, j t}$

Die IAW-Diskussionspapiere erscheinen seit September 2001. Die vollständige Liste der IAW-Diskussionspapiere von 2001 bis 2008 (Nr. 1-42) finden Sie auf der IAW-Internetseite www.iaw.edu/publikationene/iaw-diskussionspapiere.

## IAW-Diskussionspapiere seit Juli 2008:

Nr. 43
(September 2008)
Mikroökonomische Determinanten und Effekte von FDI am Beispiel Baden-Württemberg Christian Arndt / Anselm Mattes

Nr. 44
(September 2008)
The Impact of Macroeconomic Factors on Risks in the Banking Sector: A Cross-Country
Empirical Assessment
Olga Bohachova
Nr. 45
(Oktober 2008)
Effects of Dismissal Protection Legislation on Individual Employment Stability in Germany Bernhard Boockmann / Daniel Gutknecht / Susanne Steffes

## Nr. 46

(November 2008)
Trade's Impact on the Labor Share: Evidence from German and Italian Regions
Claudia M. Buch / Paola Monti / Farid Toubal

Nr. 47
(März 2009)
Network and Border Effects: Where Do Foreign Multinationals Locate in Germany?
Julia Spies

Nr. 48
(März 2009)
Stochastische Überlagerung mit Hilfe der Mischungsverteilung (Stand: 18. März 2009 - Version 49) Gerd Ronning

Nr. 49
(April 2009)
Außenwirtschaftliche Verbindungen der deutschen Bundesländer zur Republik Österreich
Anselm Mattes /Julia Spies
Nr. 50
(Juli 2009)
New Firms - Different Jobs? An Inquiry into the Quality of Employment in Start-ups and Incumbents
(Stand: 28. Juli 2009 - Version 1.3)
Andreas Koch / Jochen Späth
Nr. 51
(Juli 2009)
Poverty and Wealth Reporting of the German Government: Approach, Lessons and Critique
Christian Arndt / Jürgen Volkert
Nr. 52
(August/September 2009)
Barriers to Internationalization: Firm-Level Evidence from Germany
Christian Arndt / Claudia M. Buch / Anselm Mattes

Nr. 53
(September 2009)
IV-Schätzung eines linearen Panelmodells mit stochastisch überlagerten Betriebs-
und Unternehmensdaten
Elena Biewen / Gerd Ronning / Martin Rosemann
Nr. 54
(November 2009)
Financial Constraints and the Margins of FDI
Claudia M. Buch / Iris Kesternich / Alexander Lipponer / Monika Schnitzer

Nr. 55
(November 2009)
Offshoring and the Onshore Composition of Tasks and Skills
Sascha O. Becker / Karolina Ekholm /Marc-Andreas Muendler
Nr. 56
(November 2009)
Intensifying the Use of Benefit Sanctions - An Effective Tool to Shorten
Welfare Receipt and Speed up Transitions to Employment?
Bernhard Boockmann /Stephan L. Thomsen / Thomas Walter

## IAW-Diskussionspapiere

## Nr. 57

(November 2009)
The Responses of Taxable Income Induced by Tax Cuts - Empirical Evidence from the German Taxpayer Panel
Peter Gottfried / Daniela Witczak

## Nr. 58

(November 2009)
Reformoption Duale Einkommensteuer - Aufkommens- und Verteilungseffekte
Peter Gottfried / Daniela Witczak

## Nr. 59

(Februar 2010)
The Impact of Horizontal and Vertical FDI on Labor Demand for Different Skill Groups Anselm Mattes

## Nr. 60

(Februar 2010)
International M \& A: Evidence on Effects of Foreign Takeovers
Anselm Mattes

Nr. 61
The Impact of Regional Supply and Demand Conditions on Job Creation and Destruction (Februar 2010) Raimund Krumm / Harald Strotmann

## Nr. 62

(April 2010)
The Effects of Foreign Ownership Change on the Performance of German Multinational Firms Christian Arndt / Anselm Mattes

Nr. 63
(April 2010)
The Export Magnification Effect of Offshoring
Jörn Kleinert / Nico Zorell

Nr. 64
(März 2010)
Kundenbetreuung aus einer Hand im SGB II? - Integration versus Spezialisierung von
Fallmanagement, Vermittlung und materiellen Leistungen
Harald Strotmann / Martin Rosemann / Sabine Dann / Christine Hamacher

## Nr. 65

(Mai 2010)
The Combined Employment Effects of Minimum Wages and Labor Market Regulation A Meta-analysis
Bernhard Boockmann

Nr. 66
Remote Access - Eine Welt ohne Mikrodaten ?? (Stand: 20.06.2010, Version 18)
(Juni 2010)
Gerd Ronning / Philipp Bleninger / Jörg Drechsler / Christopher Gürke
Nr. 67
Opening Clauses in Collective Bargaining Agreements: More Flexibility to Save Jobs?
(Oktober 2010)
Tobias Brändle / Wolf Dieter Heinbach

Nr. 68
Interest Rate Policy and Supply-side Adjustment Dynamics
(Dezember 2010)
Daniel Kienzler / Kai Schmid
Nr. 69
Should Welfare Administration be Centralized or Decentralized? Evidence from
(Dezember 2010)
a Policy Experiment
Bernhard Boockmann / Stephan L. Thomsen / Thomas Walter / Christian Göbel / Martin Huber

Nr. 70
Banks in Space: Does Distance Really Affect Cross-Border-Banking?
(Februar 2011)
Katja Neugebauer
Nr. 71
An Almost Ideal Wage Database Harmonizing the ILO Database October Inquiry
(Februar 2011)


[^0]:    * We acknowledge financial support from the French and German Science Foundations (ANR and DFG)

[^1]:    ${ }^{1}$ Theoretically there are 159 different occupations, but there are three kinds of occupation 139 Government Executive Official. We handle them as three different occupations.

[^2]:    ${ }^{2}$ The ILO also collects data about hours worked. It would possible to use that information when estimating adjustment coefficients for different groups. Unfortunately, there is a huge requirement of cleaning and correcting the data, too.

[^3]:    ${ }^{3}$ We added a classification with respect to the skill level (low skilled, medium skilled, high skilled) of an occupation which is taken from the German Institute for Employment Research.

[^4]:    ${ }^{4}$ We use the imputed dataset which is described below.
    5 We use the regression equation: $W_{o, j t_{i}}=D_{o} \alpha_{o}+D_{j t} \alpha_{j t}+v_{o, j t_{i}}$, with $i=$ all countries, OECD, Non-OECD.

