INNODRIVE WORKING PAPER SERIES

Eukleed
An establishment level comprehensive data base for Germany

Bernd Görzig

January 2011

INNODRIVE Working Paper No 5.

The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement n° 214576

Eukleed
An establishment level comprehensive data base for Germany

Bernd Görzig

January 2011

Abstract

Many firm level studies rely on readily available databases as COMPUSTAT, based on published balance sheets. While bigger firms are quite reliable described in this data set, small and medium sized firms (SMEs) are not covered, bearing the danger that the conclusions are biased. To include SMEs into firm-level analysis, the Eukleed data set for Germany is created. Eukleed is a comprehensive integrated micro data set on employment, investment, and output for about 1.6 million German establishments, with around 25 million employees per year. The data set is combining three sources. The main source (SIS) is a linked employer employee data set (LEED). It supplies firm level information with respect to employment, employment characteristics, labour compensation by type of labour, and establishment characteristics. The remaining two sources are used to calibrate the firm level information with the aggregated data of the National Accounts. Although this calibration is performed for 70 industries and 16 Federal States, Eukleed has to be applied with care. In the analysis conducted here, it serves primarily as a tool for sensitivity calculations.

JEL classifications: Firm-level capital stock
Keywords: L23, D24, M10, C15

Contents

Tables

1. INTRODUCTION

Many firm level studies rely on readily available databases as COMPUSTAT, based on published balance sheets. While bigger firms are quite reliable described in this data set, small and medium sized firms (SMEs) are not covered, bearing the danger that the conclusions are biased1. To include SMEs into firm-level analysis, the Eukleed data set for Germany is created. Eukleed is a comprehensive integrated micro data set on employment, investment, and output for about 1.6 million German establishments, with around 40 million employment cases per year. The data set is combining three sources. The main source (SIS) is a linked employer employee data set (LEED). It supplies firm level2 information with respect to employment, employment characteristics, labour compensation by type of labour, and establishment characteristics. The remaining two sources are used to calibrate the firm level information with the aggregated data of the National Accounts. Although this calibration is performed for 70 industries and 16 Federal States, Eukleed has to be applied with care. In the analysis conducted here, it serves primarily as a tool for sensitivity calculations.

2. SOURCES

2.1 The German Social Insurance System (SIS) database

SIS is based on the register for all persons obliged to pay social security contributions. It supplies a nearly complete coverage of all German employees. Merely some governmental personnel and a number of low-income recipients are excluded. For each employee information is available for the day a particular job began and when it has been finished, including the income received during that period. Information is supplied upon the type of the job performed and the educational skill of the person doing the job. In addition, the location of its working place is given by an ID of the establishment where the person is working. This implies that the industry and the region where a person is working are available. An overview of SIS is given by Fritsch/Brixy (2004).

The micro data of the SIS are subject to very restrictive disclosure rules. In recent years, access has increasingly been made possible by the Research Data Centre (http://fdz.iab.de) of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB), which prepares individual datasets developed in the sphere of social security and in employment research and

[1] In addition, some authors, as McGahan/Porter (2002), drop the remaining comparative small enterprises from the COMPUSTAT data file for their analysis of the variance of profitability.
[2] If not noted otherwise, firm is used synonymously with establishment, the local production units.

makes them available for research purposes – primarily for external researchers. An overview on the current situation with respect to data availability from this source can be found in Bender/Möller (2009).

2.2 The EU KLEMS database

EU KLEMS supplies information on factors of production and output for about 70 industries (Section 9.3). It is fully integrated into the National Accounting framework of EUROSTAT. We use the EU KLEMS data files *ger_output_07I* and *ger_capital_input_07II* (http://www.euklems.net). A comprehensive description is given by Timmer/Mahony (2006). Capital input data published by EU KLEMS are available only in a 30-industry breakdown. They are extended to the 70-industry level by applying DIW calculations.

2.3 The National Accounts for the German Federal States (NA Fed)

This database is for the 16 Federal States of Germany (NA Fed 2009). It supplies information on output and factors of production by industry, fully integrated into the National Accounts for Germany. The industry breakdown with 10 industries is below the one for the EU KLEMS data base with in general 70 industries.

3. VARIABLES

3.1 Employment

EU KLEMS, as ESA 95, describes employment as the average stock of employed persons over the year. In the SIS database for each person, information is available on the period employed, measurable in days. Here, this property is called *employment case*. An *employment case* can be a person that works only for one day or it could be a person that works all the days of the year. The same person may consist of several employment cases. To make this information comparable, the employment cases are converted into individual person years, which can be summed up to the EU KLEMS industry levels.

For employment figures, no adaptations of SIS firm-level data to the EU KLEMS industry data are made. Divergences between Eukleed and EU KLEMS with respect to the industrys' employment figures are caused by three factors:

- EU KLEMS data are based on enterprises, the legal units as the smallest entity; SIS data are only available for establishments, the local units. For some industries, the number of employees in establishments is higher than in the enterprises of these industries (i.e. industry *DF*).
- SIS does not cover certain types of employees in institutional sectors S.14/S.15 with

an impact for industries *L* and *M*.

- SIS does not cover very low-income recipients (i.e.in industry *H*).

For all industries considered in the analysis, the coverage compared with EU KLEMS data is about 79% if only the figures of the industries selected for the analysis are compared. With respect to total employment, considering also the industries excluded from the analysis, the coverage is around 60%. These relations are valid within certain margins also for all other variables in the analysis.

Table 1: Coverage of employment - averages 1999 - 2003

industry description	Nace rev1	Innodrive establishment values compared with EU KLEMS enterprise values
FOOD, BEVERAGES AND TOBACCO	DA	0,74
TEXTILES, TEXTILE, LEATHER AND FOOTWEAR	DB,DC	0,84
WOOD AND OF WOOD AND CORK	DD	0,93
PULP, PAPER, PAPER, PRINTING AND PUBLISHING	DE	0,80
Coke, refined petroleum and nuclear fuel	DF	1,18
Chemicals and chemical products	DG	0,90
Rubber and plastics	DH	0,90
OTHER NON-METALLIC MINERAL	DI	0,80
BASIC METALS AND FABRICATED METAL	DJ	0,96
MACHINERY, NEC	DK	0,90
ELECTRICAL AND OPTICAL EQUIPMENT	DL	0,96
TRANSPORT EQUIPMENT	DM	0,71
MANUFACTURING NEC; RECYCLING	DN	0,90
ELECTRICITY, GAS AND WATER SUPPLY	E	0,81
CONSTRUCTION	F	0,79
WHOLESALE AND RETAIL TRADE	G	0,71
HOTELS AND RESTAURANTS	H	0,50
TRANSPORT AND STORAGE AND COMMUNICATION	I	0,91
FINANCIAL INTERMEDIATION	J	0,87
RENTING AND BUSINESS ACTIVITIES exc. Real estate	K	0,68
HEALTH AND SOCIAL WORK	N	0,79
OTHER COMMUNITY, SOCIAL AND PERSONAL SERVICES	O	0,68
All selected industries		0,79

3.2 Self-employed

It is assumed that each establishment is associated with one self-employed person, except in industries where in the EU KLEMS data base self-employment is not defined. For companies with an own legal status this assumption certainly does not hold. In general, these companies are

very big and the induced estimation error is expected to be negligible. In the database, around 160.000 establishments with a production value of more than two million Euros can be found. They cover merely 20% of all analysed establishments but nearly 70% of total employment.

No estimates are made for the number of family workers.

3.3 Wage expenditures

Wages in the SIS database do not include social security contributions completely. Furthermore, they are censored for low incomes and for high incomes depending on the region and the year considered. To be more precise, in this sample we do not have sufficient information on employees with a monthly wage below 400 Euro. The coverage for this type of employees is very low in some cases. The number of people covered has varied over time due to changes in the respective legislation.

Table 2: Coverage of wage expenditures - averages 1999 - 2003

industry description	Nace rev1	Innodrive establishment values compared with EU KLEMS enterprise values
FOOD, BEVERAGES AND TOBACCO	DA	0,74
TEXTILES, TEXTILE, LEATHER AND FOOTWEAR	DB,DC	0,84
WOOD AND OF WOOD AND CORK	DD	0,93
PULP, PAPER, PAPER, PRINTING AND PUBLISHING	DE	0,77
Coke, refined petroleum and nuclear fuel	DF	1,20
Chemicals and chemical products	DG	0,91
Rubber and plastics	DH	0,90
OTHER NON-METALLIC MINERAL	DI	0,80
BASIC METALS AND FABRICATED METAL	DJ	0,97
MACHINERY, NEC	DK	0,90
ELECTRICAL AND OPTICAL EQUIPMENT	DL	0,92
TRANSPORT EQUIPMENT	DM	0,73
MANUFACTURING NEC; RECYCLING	DN	0,90
ELECTRICITY, GAS AND WATER SUPPLY	E	0,81
CONSTRUCTION	F	0,79
WHOLESALE AND RETAIL TRADE	G	0,73
HOTELS AND RESTAURANTS	H	0,49
TRANSPORT AND STORAGE AND COMMUNICATION	I	0,91
FINANCIAL INTERMEDIATION	J	0,88
RENTING AND BUSINESS ACTIVITIES exc. Real estate	K	0,70
HEALTH AND SOCIAL WORK	N	0,79
OTHER COMMUNITY, SOCIAL AND PERSONAL SERVICES	O	0,70
All selected industries		0,79

At the upper end, the income reported is censored by limits, which are set by law for the social insurance contributions. Contrary, all other characteristics of employees in the SIS are available also for these employees. We apply a wage function to estimate all wages outside the upper wage limits given by the data set, using fixed effect regressions with about 20 different explaining variables. Since the LEED data set is very big, comprising about 140 million employment cases for the period considered, a multitude of explaining variables could be included both of the firm-specific and person-specific type.

Grossing up wages and days worked to the State and industry levels as given by NA Fed and EU KLEMS respectively, we can calculate average wages per day for each industry and State. The average wage per day is adapted to the respective value in these data sets. Multiplying daily wages for all employees in the firm by the days worked results in total wage expenditures of the firm, which is a central variable for the subsequently described estimates.

The coverage of wage expenditures by industry is very similar to the findings for employment. In the average, the coverage is about 80%.

3.4 Capital formation

Capital formation can be distinguished between tangibles and intangibles. Both types of capital formation, tangibles and intangibles, are estimated at firm level using the available information on personal (age, gender, etc.), occupational (type of occupation, tenure, mobility, etc.), and educational characteristics (school and university degree, etc.) for employees in the establishment.

3.4.1 Tangibles

According to the classification applied in the EU KLEMS database, several types of capital formation can be defined as tangibles:

Equipments are divided by type

- Transport equipment,
- Information equipment,
- Communication equipment, and
- All other equipment.

Buildings are divided into

- Dwellings,
- Infrastructure, and
- Other buildings.

In addition, EU KLEMS allows for

Other assets,

- Software, and
- Some other intangibles as quantified in the National Accounts.

In the data set for the Federal States (NA Fed), only between *buildings* and *equipments* (including *other assets*) can be distinguished.

Table 3: Coverage of tangible capital formation - averages 1999 - 2003

industry description	Nace rev1	New intangible capital formation compared with EU KLEMS values
FOOD, BEVERAGES AND TOBACCO	DA	0,74
TEXTILES, TEXTILE, LEATHER AND FOOTWEAR	DB,DC	0,85
WOOD AND OF WOOD AND CORK	DD	0,95
PULP, PAPER, PAPER, PRINTING AND PUBLISHING	DE	0,77
Coke, refined petroleum and nuclear fuel	DF	1,20
Chemicals and chemical products	DG	0,89
Rubber and plastics	DH	0,90
OTHER NON-METALLIC MINERAL	DI	0,80
BASIC METALS AND FABRICATED METAL	DJ	1,01
MACHINERY, NEC	DK	0,90
ELECTRICAL AND OPTICAL EQUIPMENT	DL	0,97
TRANSPORT EQUIPMENT	DM	0,71
MANUFACTURING NEC; RECYCLING	DN	1,02
ELECTRICITY, GAS AND WATER SUPPLY	E	0,80
CONSTRUCTION	F	0,80
WHOLESALE AND RETAIL TRADE	G	0,74
HOTELS AND RESTAURANTS	H	0,49
TRANSPORT AND STORAGE AND COMMUNICATION	I	1,03
FINANCIAL INTERMEDIATION	J	0,89
RENTING AND BUSINESS ACTIVITIES exc. Real estate	K	0,72
HEALTH AND SOCIAL WORK	N	0,82
OTHER COMMUNITY, SOCIAL AND PERSONAL SERVICES	O	0,72
All selected industries		0,81

As for intangibles, also for tangibles, it is assumed that a relation exists between certain types of occupation and the respective types of capital formation. In the case of tangibles, we use the wage information on certain types of occupations as an indicator for the investment into this type of asset. This certainly very crude assumption might be acceptable in cases, where a limitational relation between capital and labour (cars & drivers, machines & machinists) can be assumed. It is obvious that infrastructure in many cases cannot related directly to employment. Similar, in the

case of dwelling and in particular in the case of owner occupied dwellings this methodology cannot be applied (See also the health warnings by EU KLEMS 2006). A more adequate approach would be to apply information on houses and flats[3]. Industries, where infrastructure and dwellings play a major role are left out in the analysis. In addition, for the remaining industries we made sensitivity calculations in changing the assumptions on the distribution of tangible capital formation across firms by type of asset.

National Accounts as well as business accounts set a lower limit for expenditures to be counted as capital formation. We treat all firm specific expenditures below 1.000 Euro as intermediate consumption.

If firms need capital for production, they do not necessarily buy assets but rent them instead. This is particularly the case for bigger not separable types of assets in smaller firms, which play a dominant role in our data set. For instance, in the German investment survey frequently firms report that they do not have any investment at all. We assume for each type of asset separately that all expenditures below the fifth percentile in the respective EUKLEMS industries are not capital formation but intermediate consumption.

Depending on the impact of divergences between establishment and enterprises in the individual industries, the coverage of capital formation varies across industries. In the average for all selected industries, it is around 81%.

3.4.2 Intangibles

Capital formation for intangibles so far not considered in the National Accounts (new intangibles) is calculated in line with the basic idea as developed in the INNODRIVE project (Piekkola 2009). Each firm is assumed to produce goods of the types

- Information and communication (*ICT*),
- Research and development (*R&D*), and
- Organisational competencies (*OC*).

It is assumed that the production of these types of goods is exclusively directed towards own uses by the firm. If the uses are not in the current year, these types of goods can be classified as intangible capital goods, which are not counted as investment in conventional calculations of capital stocks and depreciation, as for instance in the National Accounts. In order to produce

[3] The National Accounts define investment into dwellings as an entrepreneurial activity even if it is done by Private Households (S.14). However, it would be in the logic of this treatment that also the possibly part-time labour input of Private Households for managing the flats and houses would be counted. Since this input contains a considerable amount of activity with an expected return in the future, it would also be relevant for a comprehensive measurement of intangibles.

these types of capital goods, firms apply resources supplied by different factors of production: labour, intermediate, and capital services.

To assess labour services, we separate three types of labour input, according to the type of produced goods: *ICT*-, R&*D*-, and *OC*-related personnel (see table 9.2).

- *ICT* personnel is defined by information and communication experts.
- R&*D* personnel is defined by technicians, engineers, and similar.
- *OC* personnel is defined by management (incl. owners) and marketing employees

We assume that from these types of labour input only a certain proportion is engaged in the production of new intangible goods. The remaining employees of the respective types of labour are engaged in current production.

Substantial parts of intangible *ICT* investment are software and databases. We see the possibility of double counting here, since the calculation of own account software already included in the National Accounts are partly based on a similar methodology. For *ICT*, we are assuming that half of the expenditures for labour are dedicated to the maintenance of *ICT* operation and half to the production of own account capital goods.

We apply a very broad definition of R&*D* personnel including all employees that have a technical education. Accordingly, only 70% of the labour input (measured by the share in total labour cost for R&*D* personnel) is assumed to be engaged in creating R&*D* capital goods for future uses by the firm.

Organisational workers include management and marketing people. In addition, we consider part of the labour input done by self-employed as own account organisational capital formation. Labour input of self-employed consists in wide parts in management and marketing activities. Marano/Haskel (2006) are mentioning this, but found it not clear whether this type of labour input can be treated as part of the organisational capital formation in a similar way as the one of employed managers. In our data set, the importance of self-employed is comparatively big since the majority of the establishments consist of very small units. Only 20% of labour compensation for OC-personnel is considered to contribute to the production of capital goods.

Table 4: Share of labour cost dedicated to the production of intangible goods

	ICT	R&*D*	*OC*
Share of labour input for own account production	0.50	0.70	0.20

Different from Corrado/Hulten/Sichel (2004), hereafter CHS, we also evaluate the value of intermediate and capital cost related to labour cost necessary in own account production of

intangible capital goods. We refer to those industries, which are engaged in market production of comparable type of goods. The following industries have been chosen:

- Computer and related activities (Nace 72) as proxy for *ICT* goods,
- Research and development (Nace 73) as proxy for *R&D* goods, and
- Other business activities (Nace 74) as proxy for *OC* goods.

The amount of intermediate and capital expenditures is evaluated by two factors respectively. These factors show the magnitude by which expenditures for labour have to be multiplied to account for the additional intermediate or capital cost needed to produce the capital goods. The implicit assumption is that cost shares are bases on production function with constant returns and perfect competition prevails. The two factors for intermediate and capital cost have been used to form a combined multiplier, which can be applied at firm level to generate total production cost for the different types of intangible capital goods.

Data for the assessment of the factors are taken from the EU KLEMS database for Finland, UK, Germany, Czech Republic, and Slovenia in the NACE industries 72, 73, and 74. We assume that the average relation between the production factors labour, intermediates, and capital in these industries can also be taken as an indicator for the cost structure in own account production of these types of goods in the firms in our analysis.

Compensation for labour services is defined in line with ESA95 definitions for the National Accounts including all labour cost accruing to the firm, including also social security contributions. Similarly, intermediate consumption is defined in line with ESA95 definitions for the National Accounts. The composition of intermediate input in NACE 72 to 74 differs from the one used in the own account production of intangible products. Supply and use tables (e. g. Destatis 2006) make transparent that more than half of the intermediate input in these NACE industries comes from intra-industry supply of the same industry and possibly very often the same firm. To avoid overestimating the impact of intermediate input in the production of intangibles we only consider those intermediates, which are supplied by other industries. Accordingly, the ratio for intermediate - to labour costs has been set to 35% of the industry specific value for *ICT*- and *R&D* -products and to 45% for *OC*-products.

This ratio of intermediate to labour costs varies quite substantially across countries. In particular, in *R&D* production this ratio is 0.55 in Germany and with 0.28 only half of this in the UK (see Table 5). The comparative high value in Germany for intermediates in the case of *R&D* cannot be seen as a purely statistical outlier. It is confirmed by the evidence that at least in manufacturing the additional intermediate input is around 50% of the cost for the personnel (EU-Structural business survey, e.g. DESTATIS 2003). This might have to do with the fact that *R&D* activities in Germany, in many cases have been transferred to special research companies in the ownership of manufacturing firms. This could imply that the part of purchased *R&D* (from own affiliates)

might have a bigger impact compared with own account production. However, it is likely that purchased *OC* activities from own affiliates are also important in international firms.

Table 5: Combined multiplier based on EU KLEMS averages (1999 - 2003)

Type of intangible industry proxy	*ICT* Nace 72	*R&D* Nace 73	*OC* Nace 74
Related intermediate costs[1] - Labour costs=1			
Germany	0.25	0.55	0.51
UK	0.36	0.28	0.53
Finland	0.31	0.25	0.50
Czech Republic	0.54	0.36	1.47
Slovenia	0.41	0.24	0.74
Related capital costs[2] - Labour costs=1			
Germany		0.25	
UK		0.11	
Finland		0.06	
Czech Republic		0.21	
Slovenia		0.00	
Combined factor [3]			
Germany	1.50	1.80	1.76
UK	1.47	1.40	1.64
Finland	1.37	1.31	1.56
Czech Republic	1.75	1.57	2.68
Slovenia	1.41	0.24	1.74
Weighted average[4]	1.48	1.55	1.76

[1] 35% (45%) of the relation applied in the production of Nace 72, 73, (and 74).

[2] Depreciation + net capital*0.04 (external rate of return).

[3] Total production cost (excluding profits and including labour costs) related to labour costs.

[4] GER 0.40, UK 0.30, FIN 0.15, CZ 0.075, SL 0.075

The concept of capital user cost (OECD 2001, Jorgensen 1963) has been applied to quantify the expenditures needed to use capital in the production of intangible capital goods. The main components of the capital user cost are depreciation and return on capital. According to ESA95, in the case of own account production valuation has to made at basic prices. This implies that profits should be excluded from the calculations of the capital user cost. We apply an external rate of return[4] (representing the market interest rate) of 4%. Multiplied with net capital stock, this yields the interest part of the user cost.

Unfortunately, the industry breakdown for the capital accounts in EU KLEMS is less detailed than for labour and intermediates. Only one factor for all three types of intangibles can be calculated here. Calculations that are more precise can be made for Germany, referring to the

[4] We do not include changes in the prices for investment goods, because of the lack of data there is no firm level variation in this variable.

Eukleed database (cautiously applying the relations for the 10th percentile of all establishments in the sample), which shows that capital cost for producing intangibles do not have much influence on the total factor such that their share in overall production cost is comparatively small. Note that this refers only to conventional capital usage, not including the cost for the use of intangible capital itself, which is not calculated here.

Table 6: Central settings for intangibles

	ICT	*R&D*	*OC*
Investment share of labour input	0.50	0.70	0.20
Combined factor for other inputs	1.48	1.55	1.76
Final multiplier on labour costs	**0.70**	**1.10**	**0.35**
Depreciation rate	**0.33**	**0.20**	**0.25**

We use the weighted average relationship between labour, intermediates, and capital in NACE 72-74 calculated in table 5 as a proxy for the cost structure of own account production of intangible goods in the firms in our analysis. Combined with the figures in Table 4 (replicated in Table 6) on the share of labour costs dedicated to the production of intangible capital, we arrive at a combined multiplier of 0.35 for *OC*, 1.10 for *R&D*, and 0.70 for *ICT*. The central settings on intangibles are thus shown in Table 6.

For several reasons, the average structure of the selected business services may not adequately represent the actual cost shares for producing own account capital goods by the firms. The calculations contain a number of uncertainties:

- The sampling rate in the business services is likely to vary from one country to another.
- The relations may change depending on the business cycle.
- The results for the selected industries depend heavily on the different degree of outsourcing in the countries. Capital services, for instance office space, might be bought (rented) or supplied by own investment.
- The cost structure may vary depending on whether the produced goods in the chosen industries are capital goods or other types of goods.
- The relations found are weighted averages across all firms of the industry in question.

It is evident that the combined multiplier calculated here is a very rough indicator for capturing the cost structure in the production of intangible capital goods of a firm. We therefore apply for all countries a common set of combined multipliers for each intangible type *ICT*, *R&D*, and *OC*. Multiplied with the wage expenditures of a firm, these factors yield an assessment of the investment estimates of the firms' own account production in the intangible capital categories in question. The common multiplier is calculated by weighting the country specific outcomes with the weights of the countries in the aggregated GDP of these countries.

Table 7 shows that the impact of intangible compared with tangible investment varies considerably across industries. It is very high for firms in machinery and electrical and optical equipment, but also for construction (Due to the high proportion of self-employed). Low values can be found for some service industries like transport and health. It should however be kept in mind that these figures can only give a first impression of the importance of intangibles. In the average for all firms, intangible investment amounts to 85% of conventional investment.

Table 7: Intangible capital formation by industry - averages 1999 - 2003

industry description	Nace rev1	Intangible as percentage of conventional capital formation
FOOD, BEVERAGES AND TOBACCO	DA	0,46
TEXTILES, TEXTILE, LEATHER AND FOOTWEAR	DB,DC	1,12
WOOD AND OF WOOD AND CORK	DD	1,04
PULP, PAPER, PAPER, PRINTING AND PUBLISHING	DE	0,41
Coke, refined petroleum and nuclear fuel	DF	0,58
Chemicals and chemical products	DG	1,16
Rubber and plastics	DH	0,66
OTHER NON-METALLIC MINERAL	DI	0,73
BASIC METALS AND FABRICATED METAL	DJ	0,88
MACHINERY, NEC	DK	2,20
ELECTRICAL AND OPTICAL EQUIPMENT	DL	2,18
TRANSPORT EQUIPMENT	DM	0,76
MANUFACTURING NEC; RECYCLING	DN	0,90
ELECTRICITY, GAS AND WATER SUPPLY	E	0,33
CONSTRUCTION	F	2,27
WHOLESALE AND RETAIL TRADE	G	1,50
HOTELS AND RESTAURANTS	H	0,56
TRANSPORT AND STORAGE AND COMMUNICATION	I	0,15
FINANCIAL INTERMEDIATION	J	0,86
RENTING AND BUSINESS ACTIVITIES exc. Real estate	K	1,00
HEALTH AND SOCIAL WORK	N	0,34
OTHER COMMUNITY, SOCIAL AND PERSONAL SERVICES	O	0,40
All selected industries		0,80

3.5 Capital stock

The different types of capital formation as defined by EU KLEMS are used to calculate firm-level capital stocks (Timmer/Mahony 2006). The capital stock methodology and depreciation rates from the EU KLEMS project are applied (Chapter 9.1). Capital stock at historical prices as in commercial accountancies is calculated. Using the EU KLEMS methodology, the closing stock K_t for an establishment is given by:

$$K_t = K_{t-1}(1-\delta) + I_t, \qquad (1)$$

with I_t for the capital formation of the current year and a constant depreciation rate δ. In order to be able to compare capital with labour input and production, the average annual stock is calculated as the arithmetic mean of the opening and the closing stock in a year.

For intangibles, we use the depreciation rates as applied in the INNODRIVE project: 0.33 for *IT*, 0.2 for *R&D* assets, and 0.25 for *OC* assets. The latter value is also taken for assets produced by self-employed. For organisational capital, we apply a lower rate than those applied in other studies (CHS 2006). While high depreciation rates are convenient for capital stock calculations since they are relaxing the requirements for the length of the time series needed to estimate reliable capital stock figures, there is also evidence that depreciation might be lower, or in other words that the returns of the investment holds for a longer period.

In some countries (OECD 2006), as for instance Germany, tax authorities allow for a 5 years linear depreciation period on a *company value* bought by a firm. Translated into the applied EU KLEMS methodology with geometric depreciation patterns[5], one would expect a depreciation rate for the *company value* of 0.2 or below. The *company value* in this case is defined as the difference between the amount paid for the company and the sum of the replacement cost for all assets accounted for in the balance sheet. The *company value* includes more than the own account organisational capital as in our calculations, for instance also purchased intangibles, but also tangibles, which are still in use although they have been fully depreciated. High depreciation rates for selected parts of the *company value* imply lower depreciation rates for the remainder. Furthermore, for most assets for a number of reasons, tax depreciation rates are lower than the effective economic depreciation rates.

Starting values for capital stocks are calculated by using a modified version of a methodology suggested by Griffith (1999). The relation between capital formation and capital stock by type of asset and industry calculated from the EU KLEMS database is used. This relation is applied on firms existing at the first day of our observation period (January 1st, 1999) to calculate the opening stock of firm-specific capital. Capital stock calculations are based on observed figures for investment and an estimate of the initial closing capital stock $K_{\theta-1}$ in the year before we start to observe a firm in the data. We assume a constant growth of investment g before the first year of observation. Let θ be the first observation for a firm. Back extrapolating yields:

$$I_{\theta-1} = I_\theta(1-g) \qquad (2)$$

[5] We apply simple geometric depreciation and do not refer to the BEA declining balance model, where a 5 years service live assumption could result in a .33 depreciation rate, if the balance rate is set to 1.65.

with I_t for the capital formation[6] of the current year and a constant growth rate g. Given the general cumulative definition of the closing stock in equation 2, we can apply the following equation to calculate the initial stock:

$$K_{\theta-1} = I_{\theta-1} \sum_{0}^{\infty} (1-\delta-g)^t \ . \qquad (3)$$

δ is the depreciation rate and g is the growth of investment in the years preceding the initial year. Applying the sum formula for a geometric row leads to

$$K_{\theta-1} = \hat{I} \ \frac{1-(1-\delta-g)^T}{1-(1-\delta-g)} . \qquad (4)$$

The initial investment $\hat{I}$ stands for the starting value $I_{\theta-1}$ for the back extrapolation, using the growth rate of investment g before the first observation. In theory, *T* should be infinite, for practical purposes it can be set to 100. Growth rate g depends on the average growth rate of intangible investment in the observation period. This implies that we assume that the past and current average growth rates are similar. $\hat{I}$ is set to be the average investment in the five-year period following the first observation year θ. The average is used to assess the average investment over the business cycle. It is corrected by a discount factor to reflect the growth of investment in the observation period.

3.6 Value added

Two ways to quantify value added are possible:

- Adding up its components, or
- Subtracting intermediate input from gross output.

Here, in a first step, firm specific value-added is estimated by adding up its components, which are

- Labour compensation,
- Taxes minus subsidies,
- Depreciation,
- Labour compensation for self-employed,
- Operating surplus (excl. labour compensation for self-employed).

[6] According to the definition given by ESA 95, gross fixed capital formation (GFCF) of a firm is defined as new investment plus acquisition of used assets minus the sale of used assets.

Labour compensation is described in the preceding section.

Taxes minus subsidies are estimated by applying the relation with value-added given for the EU KLEMS industries.

Depreciation can be quantified at firm level, using the capital-stock calculations described in the preceding section. Divergent from EU KLEMS it is valued at historical prices.

Labour compensation for self-employed is quantified by using firm-specific data on wages paid. It is assumed that the wage equivalent for the labour input of the self-employed is at least as high as the highest annual wage paid by the respective firm. This assumption results in an imputed value for the economy as a whole, which is considerable below the value calculated in EU KLEMS, which is based on the average wage level of employees. It should be noted that Eukleed consists of a very high number of small firms with only a small number of employees, usually having a low qualification and thus a low wage. About one third of the establishments in the analysis have less then five employed persons. A considerable number of establishments consist of a - presumably - high qualified self-employed (i.e. doctors, engineers, lawyers) and very few low qualified employees, which may also do part-time work (i.e. secretaries, office personnel).

Industry specific adaptations to the EU KLEMS database for the imputed labour cost of self-employed are made, assuming that the self-employed are also the recipients of the imputed wages for family workers.

Operating surplus (after the deduction of the labour compensation for self-employed) is estimated by multiplying in a first step industry specific rates of return, taken from the EU KLEMS dataset, with the calculated capital stock as described in the preceding section. If EU KLEMS return rates are negative, the rate of return is set to 0.01. Return rates in EU KLEMS are published only in a 30-industry breakdown. For the 70-industry breakdown, the same rates are used as in the heading industries. The average annual capital stock, the mean of opening and closing stock of the year, is applied.

In a second step, the calculated firm-specific estimates of value added are adapted to the EU KLEMS and NA Fed figures of value added. Since depreciation is determined endogenously, labour compensation is based on primary data sources, taxes minus subsidies have a low weight for the majority of industries, only operating surplus or imputed labour compensation can be changed to adapt firm-level value added. In general, firm-specific operating surplus is corrected proportionally according to the industry-specific and State-specific values given by the EU KLEMS and the NA Fed data set. In the case of negative values for operating surplus originating already in the EU KLEMS industry data, labour income of self-employed is reduced (e.g. industries *H* and *K*). Note that this does not prevent that operating surplus of individual firms becomes negative.

Table 6: Coverage of value-added - averages 1999 - 2003

industry description	Nace rev1	Innodrive establishment values compared with EU KLEMS enterprise values
FOOD, BEVERAGES AND TOBACCO	DA	0,74
TEXTILES, TEXTILE, LEATHER AND FOOTWEAR	DB,DC	0,84
WOOD AND OF WOOD AND CORK	DD	0,93
PULP, PAPER, PAPER, PRINTING AND PUBLISHING	DE	0,76
Coke, refined petroleum and nuclear fuel	DF	1,20
Chemicals and chemical products	DG	0,90
Rubber and plastics	DH	0,90
OTHER NON-METALLIC MINERAL	DI	0,80
BASIC METALS AND FABRICATED METAL	DJ	0,97
MACHINERY, NEC	DK	0,90
ELECTRICAL AND OPTICAL EQUIPMENT	DL	0,93
TRANSPORT EQUIPMENT	DM	0,72
MANUFACTURING NEC; RECYCLING	DN	0,93
ELECTRICITY, GAS AND WATER SUPPLY	E	0,81
CONSTRUCTION	F	0,79
WHOLESALE AND RETAIL TRADE	G	0,74
HOTELS AND RESTAURANTS	H	0,49
TRANSPORT AND STORAGE AND COMMUNICATION	I	1,09
FINANCIAL INTERMEDIATION	J	0,86
RENTING AND BUSINESS ACTIVITIES exc. Real estate	K	0,68
HEALTH AND SOCIAL WORK	N	0,79
OTHER COMMUNITY, SOCIAL AND PERSONAL SERVICES	O	0,66
All selected industries		0,79

Table 8 shows that the for most industries the establishment values for value added are quite close to enterprise values as calculated by EU KLEMS. In the average, where the difference between the enterprise and the establishment concept is negligible, the coverage of value-added is 80%.

4. DEFLATORS

In general, all calculations are made at current prices. Based on the EU KLEMS database, industry-specific deflators could be applied to calculate volumes. However, care must be taken in the assessment of these volume data. They might be acceptable for capital formation and capital stock at firm level, since investment prices might not differ considerably between investing firms. Calculated volume indicators for value-added and gross production might be suited to correct for industry specific inflation, but cannot be used to analyse the firm-specific price-setting behaviour. Intangibles are deflated with the price indices for software.

5. UNITS

Eukleed, like the SIS as the main source for the firm data, is based on establishments. EU KLEMS and NA Fed data are organised according to the enterprise concept. Differences between both concepts arise mainly in the case of bigger companies, where one enterprise may consist of several establishments. In Eukleed, it is assumed that the key relations as described in the National Accounts, as wages per day and labour productivity are the same for establishments and industries in a given industry or region.

Only establishments with at least one employee (employment year) that remains in existence at least for two years are considered. All establishments can be analysed as panel data. From the 1.6 million establishments, roughly 1 million exist for the whole period of five years.

6. INDUSTRIES

The adaptation procedure is applied for all EU KLEMS industries. For some industries, the results should be treated very careful. In the analysis, the following industries are not included: *A* (Agriculture, Forestry), *B* (Fishing), *C* (Mining and Quarrying), and NACE *70* (Real estate), *M* (Education), *L* (Public administration, Social security), and *P* (Private Households). Reasons for excluding these industries are either their insufficient coverage in our database (*A*, *B*, and *P*) or their treatment in the National Accounts (*C*, *70*, *L*, *M*, and *P*) or both. A better understanding, why these industries have been excluded is given by EU KLEMS health warnings in EU KLEMS (2006) referring to industries: *A* to *C*, *L* to *N*, *60* to *63*, *70*, and *P*.

The exclusion of these industries reduces the number of establishments included in the analysis by 100.000 observations to remaining 1.5 million. The wage sum covered accounts for 90% of the wage sum calculated for all establishments in the SIS database.

7. OBSERVATION PERIOD

Eukleed is a true panel. It covers all days between 1999 and 2003. The period begins at the first day in January 1999 and ends the last day in December 2003.

8. REFERENCES

Bender, Stefan, Joachim Möller (2009). Labor Market I, Data from the German Federal Employment Services, Working Paper Series of the Council for Social and Economic Data 96.

Corrado, Carol; Charles Hulten, Daniel Sichel (2004). Measuring Capital and Technology: An Expanded Framework. Finance and Economic Discussion Series. Federal Reserve Board. 2004-65.

Corrado, Carol; Charles Hulten, Daniel Sichel (2006). Intangible Capital and Economic Growth. NBER working paper No. 11948.

DESTATIS (2003). Kostenstruktur der Unternehmen des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung der Steine und Erden. Statistisches Bundesamt: Fachserie 4, Reihe 4.3.

DESTATIS (2004). Beschäftigte, Umsatz und Investitionen in den Unternehmen und Betrieben des Bergbaus und des Verarbeitenden Gewerbes. Statistisches Bundesamt: Fachserie 4, Reihe 4.2.1.

ESA 1995, European system of accounts.

EU KLEMS (2007). GROWTH AND PRODUCTIVITY ACCOUNTS, Version 1.0 PART I Methodology. Prepared by Marcel Timmer, Ton van Moergastel, Edwin Stuivenwold, Gerard Ypma and Mary O'Mahony and Mari Kangasniemi.

Fritsch, Michael, Udo Brixy (2004). The Establishment File of the German Social Insurance Statistics. Journal of Applied Social Science Studies 124: 1 183-190.

Griffith, Rachel (1999). Using the ARD establishment level data to look at foreign ownership and productivity in the United Kingdom. The Economic Journal 109: F416-F442.

Hao, Janet Xiaohui, Vlad Manole (2008). Intangible Capital and Growth – an International Comparison. Paper Prepared for the 30th General Conference of The International Association for Research in Income and Wealth.

IAB (2008). BA-Beschäftigtenpanel 1998-2006. Anlageband: Codebuch.

Jorgenson, D. W. (1963). Capital Theory and Investment Behaviour, American Economic Review, 53(2), pp. 247-259.

Marrano M. G., J. Haskel, (2006). How Much Does the UK invest in Intangible Assets? Queen Mary, University of London, Working paper 578.

Marrano M. G., J. Haskel, G. Wallis (2007). What Happened to the Knowledge Economy? ICT, Intangible Investment and Britain's Productivity Record Revisited. Queen Mary, University of London, Working paper 603.

McGahan, Anita M., Michael E. Porter (2002). What Do We Know About Variance in Accounting Profitability? Management Science 48: 7 834-851.

NA Fed (2009). Bruttoinlandsprodukt, Bruttowertschöpfung in den Ländern und Ost-West-Großraumregionen Deutschlands 1991 bis 2008 Reihe 1, Band 1. Arbeitskreis "Volkswirtschaftliche Gesamtrechnungen der Länder.

OECD (2001), Measuring Capital, OECD Manual, Measurement of Capital Stocks, Consumption of Fixed Capital and Capital Services, OECD, 2001.

OECD (2006), Tax Treatment of Business Investments in Intellectual Assets: An International Comparison, Directorate for Science, Technology and Industry, STI Working Paper 2006/4.

Piekkola, Hannu (2009). Intangibles: Can They Explain the Unexplained. Department of Economics, University of Vaasa Working Papers 13.

Timmer, M., Mary O'Mahony (2006). Productivity Measurement, National Accounts Data and the System of National Accounts: Lessons from the EUKLEMS-project. Paper Prepared for the 29th General Conference of The International Association for Research in Income and Wealth.

Timmer, Marcel and Bart van Ark (2005), "IT in the European Union: A driver of productivity divergence?", Oxford Economic Papers 57: 4 693-716.

9. ANNEX

9.1 EU KLEMS Depreciation Rates

type of asset	abreviation	minimum rate	maximum rate
Residential structures	Rstruc	0.011	0.011
Non-residential structures	NRStruc	0.023	0.069
Infrastructure	Infra	0.023	0.069
Transport equipment	TraEq	0.061	0.246
Computing equipment	ICT	0.315	0.315
Communications equipment	CT	0.115	0.115
Other machinery and equipment	OMach	0.073	0.164
Products of agriculture and fores	Agri	0.073	0.164
Other products	Oth	0.073	0.164
Software and other intangibles	Soft&Int	0.315	0.315
Note: for rates by industry, see Appendix Table 1 in EU KLEMS 2007.			

9.2 INNODRIVE Classification of Intangibles

BKdl88[1]	description[2]	Characteristics of employees creating intangible assets of type:			
		ICT	R&D	Manage-ment	Marke-ting
31-32	Agricultural engineers and administrators, a.s.			All	
601-612	Engineers, physicist, mathematicians, a.s.		Low	High	
621-635	Technicians, a.s.		All		
681	Wholesale, retail trade agents, puchasing agents, a.s.			High	Low
682-688	Sales assistents, a.s.				High
691-692	Banker, a.s.			High	
703	Advertising specialists, a.s.				High
733-734	Communication experts, a.s.	All			
751-763	Chief executives, consultants, tax adviser, a.s.			All	
771-773	Financial officers, chief accountants, a.s.			High	
774	IT experts, a.s.	All			
781-782	Office executives, a.s.			High	
783	IT assistents, a.s.	All			
784-794	Office clerks, a.s.			High	
862-863	Chief executives, consultants of social institutions, a.s.			High	
881	Economists, statisticians, a.s.		All		
883	Natural scientists, a.s.		All		
911	Directors of hotels, restaurants, a.s.			High	
921	Home ecomy administrators, a.s.			High	

[1] German classification of occupations (IAB 2008; chapter 5). - [2] Translated from German - All: All employees; High: Employees with higher education; Low: Employees without higher education. - Higher education: University degree or similar (Code numbers 4 to 6 in IAB (2008; chapter 8). - a.s.: and similar. - *Sources:* IAB (2008), Piekkola (2009), own definitions.

9.3 Classification of Industries

description	Nace Rev 1	EU KLEMS	NA FED
TOTAL MANUFACTURING	D	D	D
FOOD, BEVERAGES AND TOBACCO	DA	15t16	
Food and beverages	15	15	
Tobacco	16	16	
Textiles and textile	DB	17t18	
Textiles	17	17	
Wearing apparel, dressing and dying of fur	18	18	
Leather, leather and footwear	DC	19	
WOOD AND OF WOOD AND CORK	DD	20	
PULP, PAPER, PAPER , PRINTING AND PUBLISHING	DE	21t22	
Pulp, paper and paper	21	21	
Printing and reproduction 1	220	22x	
Publishing	221	221	
Printing and reproduction 2	222	22x	
Coke, refined petroleum and nuclear fuel	DF	23	
Chemicals and chemical products	DG	24	
Chemicals excluding pharmaceuticals 1	240	24x	
Pharmaceuticals	244	244	
Chemicals excluding pharmaceuticals 2	245	24x	
Rubber and plastics	DH	25	
OTHER NON-METALLIC MINERAL	DI	26	
BASIC METALS AND FABRICATED METAL	DJ	27t28	
Basic metals	27	27	
Fabricated metal	28	28	
MACHINERY, NEC	DK	29	
ELECTRICAL AND OPTICAL EQUIPMENT	DL	30t33	
Office, accounting and computing machinery	30	30	
Other electrical machinery and apparatus nec 1	310	31x	
Insulated wire	313	313	
Other electrical machinery and apparatus nec 2	314	31x	
Electronic valves and tubes	321	321	
Telecommunication equipment	322	322	
Radio and television receivers	323	323	
Scientific instruments	331	331t3	
Other instruments	334	334t5	
TRANSPORT EQUIPMENT	DM	34t35	
Motor vehicles, trailers and semi-trailers	34	34	
Railroad equipment and transport equipment nec	350	350	
Building and repairing of ships and boats 1	351	35x	
Aircraft and spacecraft	353	353	
Building and repairing of ships and boats 2	354	35x	
MANUFACTURING NEC; RECYCLING	DN	36t37	
Manufacturing nec	36	36	
Recycling	37	37	
ELECTRICITY, GAS AND WATER SUPPLY	E	E	E
Electricity supply	400	40x	
Gas supply	402	402	
WATER SUPPLY	41	41	
CONSTRUCTION	F	F	F
WHOLESALE AND RETAIL TRADE	G	G	G
Sale, maintenance and repair of motor vehicles and motorcycles	50	50	
Wholesale trade and commission trade, except of motor vehicles	51	51	
Retail trade, except of motor vehicles and motorcycles; repair of	52	52	
HOTELS AND RESTAURANTS	H	H	H
TRANSPORT AND STORAGE AND COMMUNICATION	I	I	I
Other Inland transport	60	60	
Other Water transport	61	61	
Other Air transport	62	62	
Other Supporting and auxiliary transport activities; activities of tra	63	63	
POST AND TELECOMMUNICATIONS	64	64	
FINANCIAL INTERMEDIATION	J	J	J
Financial intermediation, except insurance and pension funding	65	65	
Insurance and pension funding, except compulsory social securi	66	66	
Activities related to financial intermediation	67	67	
RENTING AND BUSINESS ACTIVITIES	K1	K1	K1
Renting of machinery and equipment	71	71	
Computer and related activities	72	72	
Research and development	73	73	
Legal, technical and advertising	741	741t4	
Other business activities, nec	745	745t8	
HEALTH AND SOCIAL WORK	N	N	N
OTHER COMMUNITY, SOCIAL AND PERSONAL SERVICES	O	O	O
Sewage and refuse disposal, sanitation and similar activities	90	90	
Activities of membership organizations nec	91	91	
Media activities	921	921t2	
Other recreational activites	923	923t7	
Other service activities	93	93	

1 Nace rev1 industries A to C, L, M and Real estate excluded for this analysis. - *Sources:* ESA 95, EU KLEMS 2007, NA FED 2009.