



Documentation on the Register-based Back Calculation of EU-SILC 2008-2011 in Austria

December 2014

Requests

Richard Heuberger
STATISTIK AUSTRIA
Bundesanstalt Statistik Österreich
Guglgasse 13
1110 Wien
AUSTRIA
Phone: +43 (1) 711 28-8285
e-mail: richard.heuberger@statistik.gv.at

Editor and Producer

STATISTICS AUSTRIA
Bundesanstalt Statistik Österreich
Guglgasse 13
1110 Wien
AUSTRIA

Team:

Nadja Lamei (project leader)
Stefan Angel
Thomas Glaser
Susanne Göttlinger
Richard Heuberger
Elisabeth Kafka
Anneliese Oismüller
Magdalena Skina-Tabue

The product and all material contained therein are protected by copyright with all rights reserved by the Bundesanstalt Statistik Österreich (STATISTICS AUSTRIA). If the contained material is accurately reproduced and the source "STATISTICS AUSTRIA" is quoted it is permitted to reproduce, distribute, make publicly available and process the content. If data content such as tables, graphics or text published by STATISTICS AUSTRIA is partially used, displayed or otherwise changed, a note must be added at an adequate position to show data was extracted or adapted.

© STATISTICS AUSTRIA

Statistics Austria and all contributors to this publication researched and compiled the contents of this publication with due care. However, errors cannot be entirely excluded. Statistics Austria and all contributors cannot guarantee the correctness, accuracy and completeness of the contents and therefore cannot assume liability for any type of loss or damage that results, directly or indirectly, from the use of the contents presented. Please forward any corrections to the editorial team.

Vienna 2014

Table of Contents

1. Introduction	4
2. National Regulations for EU-SILC in Austria	5
3. Register data for EU-SILC in Austria: data sources	6
4. Using register data for income variables instead of questionnaire data: methodological aspects.....	8
4.1 Record Linkage	8
4.2 Income Measurement.....	8
4.3 Weighting	9
5. Differences between questionnaire data and register data 2008-2011: impact on target variables	10
5.1 Differences in household income	10
5.2 Differences in income components	12
5.3 Differences in the composition of the total household income	23
5.4 Differences in EU 2020 indicators	24
6. Conclusions	28
7. References	28

1. INTRODUCTION

Up to and including EU-SILC 2011 in Austria, income components were exclusively collected by using questionnaire tools with voluntary participation. Following the 'Income and living conditions statistics regulation (ELStV)' in Austria which came into effect in 2010, Statistics Austria is encouraged to fill income variables for EU-SILC from existing administrative records in order to decrease respondent burden and to increase data quality. Data management procedures and comparability with survey data were extensively tested in 2010. Administrative data¹ sources were then gradually introduced in EU-SILC 2011 for selected variables (Statistik Austria 2013). Since EU-SILC 2012 Statistics Austria has been using register data – if available – instead of questionnaire data on a full scale both for the calculation of household income components and for the calculation of survey weights.

However, as data production differs in various aspects depending on whether register data or questionnaire data are used this created a break in time series from 2012 onwards. Comparability issues mainly arise for income target variables but are also related to other outcomes because of differences in the calculation of weights. Changing the data source thus also affects the evaluation of the development of income-based EU 2020 indicators over time which has to be addressed. As a consequence, Statistics Austria firstly published a report in 2013 that focused on the differences between outcomes based on questionnaire data and outcomes based on administrative data for EU-SILC 2011. This included descriptive statistics for distributions of income variables and poverty risks. Based on these comparisons for 2011 it was also estimated how the usage of administrative data might change the questionnaire-based results for the equivalised income and poverty indicators in 2008-2010. These regression-based estimations provided some benchmarks for the total population. Estimations for sub-groups and for the effect on survey weights (and thus non-monetary variables) for 2008-2010, on the other hand, turned out to be unsatisfactory. Thus, the Federal Ministry for Labour, Social Affairs and Consumer Protection together with Statistics Austria decided to back-calculate all data for 2008-2010 based on administrative registers. This allows for shifting back the break in time series from between 2011 and 2012 to between 2007 and 2008 and provides comparable datasets for the evaluation of the EU 2020 social inclusion indicators that use EU-SILC 2008 as baseline.

This report deals with the integration of register data into EU-SILC. Changes in terms of methodology and results are discussed and compared to the hitherto published figures based on questionnaire data for 2008-2011. In the following two chapters both the legal context of EU-SILC in Austria as well as national administrative data sources used for calculating income variables and weights are described. The fourth chapter deals with methodological consequences of substituting questionnaire data with register data. It comprises a brief description of how sample persons are linked with administrative records. Afterwards, pros and cons of both data sources for income measurement are discussed, and the weighting process is explained. Chapter 5 contains a comparison of the effects of the revised methodology on the amount and distribution of household income, its components and EU 2020 indicators. The report concludes (Chapter 6) with a short summary and some recommendations for data users.

¹ The terms "administrative data" and "register data" are used synonymously in this report without regard of the quality of the data source. They refer to data not derived from a questionnaire. In the tables the abbreviation RD is used for "register data".

2. NATIONAL REGULATIONS FOR EU-SILC IN AUSTRIA

The European Community Statistics on Income and Living Conditions (EU-SILC) is an important source for comparative statistics on income distribution and social exclusion at the European level and also the main source for social monitoring in Austria. The collection of data for EU-SILC in the Member States is based on EU regulations. In Austria, since 2010, there also exists the additional national regulation labelled '*Einkommens- und Lebensbedingungen-Statistikverordnung - ELStV*' issued by the Federal Ministry of Labour, Social Affairs and Consumer Protection (BMAK). This provides the legal basis for using administrative data in combination with a sector-specific personal identifier ('bPK'). The bPK allows an anonymized collection of statistical and administrative data of individuals (2010 ELStV §6). Selected households for the sample of EU-SILC have to be informed in writing on the application of the regulation before the first wave of the survey starts. The ELStV also states that the main goals of using administrative data are quality assurance and a reduction of respondents' burden. Funding of EU-SILC in Austria by Eurostat ended in 2007. Since 2008, EU-SILC has been solely financed by the BMAK.

Data for EU-SILC are collected for a representative sample of private households. In 2013, 5,977 households comprising 13,250 people were interviewed. The survey is conducted by means of personal interviews (CAPI or CATI) with all household members. Sampling frame for EU-SILC is the central population register ('ZMR') in Austria. Persons in collective households and persons that do not have a fixed place of residence are thus not part of the sample. In Austria, EU-SILC was first conducted in 2003 as a one-time cross-sectional survey by Statistics Austria. Since 2004, EU-SILC is carried out as an integrated longitudinal and cross-sectional survey - that is, around three-quarters of households are interviewed again in the subsequent year. The remaining quarter of the sample is refreshed every year. In 2007, the four-year-integrated cross-sectional and longitudinal survey was fully implemented for the first time and has been continued since then.

Results of EU-SILC in Austria are published annually in a report and a volume of tables for the main indicators. Furthermore, members of the SILC project team occasionally publish articles on selected topics in the 'Statistical News' (*Statistische Nachrichten*) edited by Statistics Austria and in other journals and publications. Anonymized microdata are made available for researchers free of charge.

3. REGISTER DATA FOR EU-SILC IN AUSTRIA: DATA SOURCES

The computation of household income in EU-SILC is based on recommendations of the Canberra Group, an international experts' group that developed guidelines for comparable household income statistics in 2001 and updated them in 2011 (The Canberra Group 2001, 2011). Total household income is calculated as the sum of earned income, capital gains, pensions and social transfers minus taxes and social security contributions. In the final step a household's disposable income is calculated by subtracting and/or adding (paid/received) alimonies and other private transfers between households (see also Table 1). Equivalised household income is defined as a household's disposable income divided by the sum of consumption equivalents of that household² using the so-called EU-scale (modified OECD scale): For each household the first adult receives a demand weight of 1, each additional adult gets a weight of 0.5 whereas each additional child under 14 years receives a weight of 0.3. For instance, a household with 2 adults and one child would have a total consumption equivalent of 1.8 compared to a single-person household.

Table 1 provides an overview of the data source for each target variable. Variables that are highlighted in white are based on questionnaire data, those in grey are based on register data³. As evident from that table, not all of the income components can be filled from register data. This is either due to the unavailability of particular register data sources for a variable or due to methodological reasons (i.e. the time lag for receiving final data for PY050 is too long). Table 2 contains the share of all income components on the total household income. The share of the sum of all components coming from register data on the total sum varies between 86.4% (2011) und 87.5 % (2010).

Table 1: Calculation of the total household income in EU-SILC 2008 – 2011

		Sum, billion €			
		2008	2009	2010	2011
	PY010 Employee Cash or Near Cash Income	69.918	73.585	75.727	74.221
+	PY050 Cash Benefits or Losses from Self-Employment	10.181	11.162	11.04	12.056
+	PY090 Unemployment Benefits	2.372	2.395	3.291	3.291
+	PY100 Old-Age Benefits	26.153	27.88	30.14	28.617
+	PY110 Survivor' Benefits	0.389	0.531	0.568	3.163
+	PY120 Sickness Benefits	0.535	0.563	0.538	0.715
+	PY130 Disability Benefits	2.327	2.778	2.932	2.913
+	PY140 Education-Related Allowances	0.283	0.305	0.303	0.328
+	PY080 Pension from Individual Private Plans	0.214	0.133	0.144	0.249
=	Sum of Personal Incomes	112.372	119.333	124.683	125.553
+	HY040 Income from Rental of a Property or Land	1.975	1.509	2.062	2.135
+	HY050 Family/Children-Related Allowances	5.506	6.034	5.782	6.569
+	HY060 Social Exclusion Benefits not elsewhere classified	0.249	0.231	0.356	0.34
+	HY070 Housing Allowances	0.304	0.266	0.314	0.367
+	HY080 Regular Inter-Household Cash Transfer received	1.233	1.154	1.227	1.411
+	HY090 Interest, Dividends, Profit from Capital Investments in Unincorporated Business	1.862	1.660	1.740	1.904
+	HY110 Income received by People aged under 16	0.218	0.196	0.103	0.18
=	Sum of Household Incomes	11.347	11.051	11.583	12.908
-	HY130 Regular Inter-Household Cash Transfer paid	1.363	1.410	1.564	1.549
-	HY145 Repayments/Receipts for Tax Adjustment	-0.788	-0.842	-0.787	-0.881
=	HY020 Total Disposable Household Income	123.144	129.816	135.488	137.793

Source: Statistics Austria, EU-SILC 2008-2011. Shaded income components are based on register data. Weighted results.

² This rests on the assumption that with increasing household size and depending on the age of children in the household cost savings can be achieved through joint economies.

³ However, due to register data limitations for some income components like family/children-related allowances (HY050) both questionnaire and register data had to be used.

Table 2: %-share of income components on total disposable household income

	2008	2009	2010	2011
PY010 Employee Cash or Near Cash Income	56.8	56.7	55.9	53.9
+ PY050 Cash Benefits or Losses from Self-Employment	8.3	8.6	8.1	8.7
+ PY090 Unemployment Benefits	1.9	1.8	2.4	2.4
+ PY100 Old-Age Benefits	21.2	21.5	22.2	20.8
+ PY110 Survivor' Benefits	0.3	0.4	0.4	2.3
+ PY120 Sickness Benefits	0.4	0.4	0.4	0.5
+ PY130 Disability Benefits	1.9	2.1	2.2	2.1
+ PY140 Education-Related Allowances	0.2	0.2	0.2	0.2
+ PY080 Pension from Individual Private Plans	0.2	0.1	0.1	0.2
= Sum of Personal Incomes	91.3	91.9	92	91.1
+ HY040 Income from Rental of a Property or Land	1.6	1.2	1.5	1.5
+ HY050 Family/Children-Related Allowances	4.5	4.6	4.3	4.8
+ HY060 Social Exclusion Benefits not elsewhere classified	0.2	0.2	0.3	0.2
+ HY070 Housing Allowances	0.2	0.2	0.2	0.3
+ HY080 Regular Inter-Household Cash Transfer received	1.5	1.3	1.3	1.4
+ HY090 Interest, Dividends, Profit from Capital Investments in Unincorporated Business	1	0.9	0.9	1
+ HY110 Income received by People aged under 16	0.2	0.2	0.1	0.1
= Sum of Household Incomes	9.2	8.5	8.5	9.4
- HY130 Regular Inter-Household Cash Transfer paid	1.1	1.1	1.2	1.1
- HY145 Repayments/Receipts for Tax Adjustment	-0.6	-0.6	-0.6	-0.6
= HY020 Total Disposable Household Income	100	100	100	100

Source: Statistics Austria, EU-SILC 2008-2011. Shaded income components are based on register data. Weighted results.

Based on the national regulation '*Einkommens- und Lebensbedingungen-Statistikverordnung - ELStV*' the sector-specific personal identifier (bPK) allows an anonymized linking of survey and administrative data for the sampled individuals (2010 ELStV §6). The following list contains a description of those national registers that were used for collecting income variables.

- *Dataset on HV-Qualifications*: contains entries from Austria's public social security register. This record does not include income information, but a number of variables concerning the social security status or changes in social security status for a given year, e.g. whether somebody received unemployment benefits, disability benefits etc. (2008-2011)
- *Wage Tax Dataset*: contains information on all taxable earnings; mainly employee earnings, pension income (retirement benefits), paid maternity leave (8 weeks before and after expected birth date) and sickness benefits. The data set contains information on gross income, paid social security contributions and payroll taxes paid. It thus allows the calculation of the respective net earnings. Moreover, the wage tax dataset also comprises care allowances, although they are not taxable in Austria. In sum, the wage tax dataset delivers roughly 75% of the total household income in EU-SILC. (2008-2011)
- *Pension Dataset*: data on old-age benefits received from the public pension system. This dataset is also a source for income target variables other than PY100. (2008 – 2011)
- *Tax Adjustment Dataset*: total repayments or receipts for tax adjustment in a given year for employee earnings (irrespective for how many years the tax was adjusted). (2008 – 2011)
- *Transfers Dataset*: data on unemployment benefits (on a daily basis); the dataset also contains the beginning and ending date of unemployment spells (= beginning and terminating of benefits) (2008 – 2011)
- *Dataset on Family Allowances*: information on family/children-related allowances. A differentiation with respect to the number of children or other characteristics of the children (disabilities) is not possible. One total sum for each bPK-AS is provided. (2008 – 2011)
- *Dataset for benefits to accident victims and surviving dependents*: (2010 and 2011)

4. USING REGISTER DATA FOR INCOME VARIABLES INSTEAD OF QUESTIONNAIRE DATA: METHODOLOGICAL ASPECTS

Structural differences between questionnaire and administrative data can occur both in terms of the number of income recipients as well as in terms of the income distribution. In this chapter methodological aspects of using administrative data instead of questionnaire data will be discussed briefly. The impact on numerical results for EU-SILC 2008-2011 for several target variables and subgroups of the population will be presented in chapter 5.

4.1 RECORD LINKAGE

In order to identify people ("cases") in the administrative datasets, the bPK-AS of people covered by the survey must be known. Usually this is straight-forward as the bPK-AS of individuals living in the sampled households are already collected as part of the sampling procedure. However, there always remain household members that are part of the target population (actual household members) but are not covered by the sampling frame (mainly because they are not officially registered at a particular household's address). For these persons the bpk-AS is missing ex ante. According to the ELStV identifiers for these cases, however, can be obtained from the Federal Ministry for the Interior. This was also necessary for the back calculation for 2008-2011. The share of bPK-AS over the total population in the survey that was found varies over time but generally decreases the farther back the survey year is (Table 3).

During the data editing process the linkage with the bPK-AS file revealed 5 duplicate records in the files of Persons in EU-SILC 2009 and 10 in the corresponding files in EU-SILC 2008. Additionally, a returnee which should not have been followed up in 2009 was found. This detection was possible because the bPK-AS is a unique identifier for individuals. The detected cases were excluded.

Table 3: Share of bPK-AS found for 2008 - 2011 as a percentage of all individuals in the sample

2008	2009	2010	2011
95.6	97.7	96.9	99.4

Source: Statistics Austria, EU-SILC 2008-2011.

Missing bPK-AS most often occurred among younger people, persons living in Vienna and persons with non-Austria citizenship. Particularly for people below the age of 40, it was cumbersome to link survey data with register data. With the exception of EU-SILC 2011, the proportion of missing links for women was higher than for men among all age groups (by about 1 percentage point). As a consequence, using register data results in an under-reporting of income data for those groups. Most households with a reported household income of 0 € receive an imputed income during the data editing process.⁴ However, for unlinked individuals in households with linkable people the underreporting remains, resulting in household incomes that are too low, on average.

4.2 INCOME MEASUREMENT

Until (and including) EU-SILC 2011 only questionnaires were used as survey instruments for the data collection of income variables. Using questionnaires has several consequences: measurement of income variables is potentially influenced by conditions of the surveying situation. For instance, because of social desirability and interviewer effects reported income variables may have a central tendency. Furthermore, as the reference period for income variables in EU-SILC is the preceding calendar year memory gaps for incomes that were received for a very short time period (few days) and/or rounding of small incomes seems likely. The same might be true for several non-revolving income types such as bonuses or special payments.

⁴ This applies to households in which it is not possible at all to link any person to administrative data via the bPK-AS as well as for persons without any employee income, pension income or unemployment income based on the monthly main activity status.

Thus, both very high and very low incomes may not be recalled correctly and/or remain unreported if questionnaires are used.

After 2011, register data - if available - have entirely substituted questionnaire instruments for filling income variables. Through the use of administrative data it was expected to reduce such forms of measurement error and to reduce item non-response for income variables. In general, it can be assumed that for a majority of the persons sampled data from registers are more complete and more accurate as compared to questionnaires. This particularly applies to incomes at the top and the bottom of the distribution.⁵ Thus, the income distribution is expected to be more unequal if administrative data is used. Further advantages of using administrative data (and conversely the disadvantages of questionnaire data) are the relief of respondent burden, cost savings, quicker data availability and higher accuracy because of more complete and consistent data. Specific challenges of using administrative data in statistical surveys arise from the cost of procurement, anonymisation and data linking. Furthermore, there can also be problems concerning availability or timeliness.

4.3 WEIGHTING

This section describes conceptual changes in the weighting of the cross-sections of EU-SILC 2008 to 2011 due to the use of administrative data. Register information for marginal distributions was used for the calibration of weights since the beginning of the integrated rotational design.⁶ The presence of more comprehensive administrative data in the EU-SILC sample now makes it possible to use new marginal distributions for the bounded extrapolation. Furthermore, in the calibration of marginal distributions from administrative data the same data base can now be used in the sample and the administrative data.

As an additional marginal distribution the wage tax income was used in calibration. More specifically, the household weights were calibrated to the total number of persons receiving income from employment (aged 15+ years) or from pensions. These features are available for the sampling frame and thus also for the EU-SILC sample. After extensive tests and comparisons, it was decided to calibrate the weights on the number of recipients of income from employment and pensions. Without this adjustment, there was a slight underestimation (up to 3%) in the number of income payments.

In order to avoid false representation of persons receiving unemployment benefits, which is a group highly affected by risk of poverty or social exclusion, weights were calibrated to the number of persons receiving unemployment benefits recorded in the register of Austrian Social Security Institutions (Hauptverband der österreichischen Sozialversicherungsträger). This particular marginal distribution has been part of the calibration procedure even before the use of register data for income measurement. However, the characteristic that was calibrated on (recipients of unemployment benefits for more than one month during the income reference period) then was recorded in the questionnaire. This means it could only be measured with the precision of one month and also measurement errors typical to questionnaires (e.g. social desirability) were unavoidable. Introducing the register of Austrian Social Security Institutions also in the income collection process allows for more precise measurement and guarantees that the number of recipients of unemployment benefits now has the same source in the sample and in the marginal distribution that the weights are calibrated on. The characteristic “number of persons receiving unemployment benefits” now refers to period of at least 60 days during the income reference year.

The abovementioned changes to the weighting procedure allow for better consistency of the results compared to the income tax register and - because of the fact that register data are now used both in the sample and in marginal distributions for calibration - are also better able to even out selective nonresponse of certain groups.

⁵ However, it is still unlikely that measurement error can be reduced to zero by using register data.

⁶ See Glaser/Till 2010.

5. DIFFERENCES BETWEEN QUESTIONNAIRE DATA AND REGISTER DATA 2008-2011: IMPACT ON TARGET VARIABLES

Using register data mainly affects income variables and income-based indicators. However, as mentioned before, using register data also impacts on non-income indicators because of changes in the survey weights. The aim of this section is to compare EU-SILC results based on questionnaire data with results based on register data. The focus is on total income, income components, structure of the total income and on EU 2020 indicators.

5.1 DIFFERENCES IN HOUSEHOLD INCOME

Using register data affects both the structure and the distribution of the household income as well as equivalised income. In general, the median of the total disposable household income based on register data (RD) is slightly higher as compared to questionnaire data. The increase lies between 1.81% for EU-SILC 2010 and 0.51% for EU-SILC 2011 (Table 4).

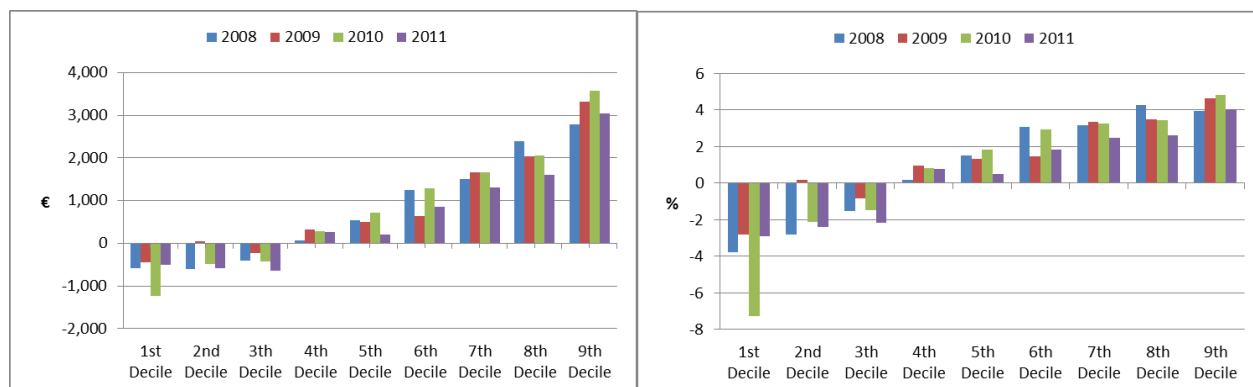
Figure 1 shows absolute and relative deviations between register and questionnaire data for each decile. It is evident, that relative deviations vary over time but are generally higher at the top and bottom of the distribution than at the middle of the distribution. For top deciles we observe a steeper distribution (Figure 1). In other words, as a consequence of using register data lower household incomes generally are lower and higher incomes are higher as compared to using questionnaire data. The total income distribution becomes more unequal which is reflected in higher values for p80p20 and p90p10. For instance, in 2008, households at the edge of the richest 10% have 4.52/4.88 times as much income as households that marginally belong to the poorest 10%. The comparison for the equivalised income results in very similar outcomes: if register data is used, the income distribution turns out to be more unequal (Table 5). This particularly applies to the bottom of the distribution which is relevant for calculating the at-risk-of-poverty rate (Figure 2).

Table 4: Total disposable household income 2008 – 2011: questionnaire and register data

	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Count (persons)	8,241,523	8,241,523	8,262,101	8,262,101	8,283,237	8,283,237	8,315,881	8,315,881
Mean	41,025	41,711	42,171	43,405	44,309	44,754	44,908	45,545
Median	35,780	36,321	37,818	38,316	39,014	39,720	40,009	40,215
p80p20	2.56	2.75	2.62	2.71	2.56	2.71	2.55	2.68
p90p10	4.52	4.88	4.41	4.75	4.36	4.92	4.34	4.65
Percentile								
10%	15,605	15,018	16,167	15,714	16,954	15,723	17,407	16,904
20%	21,803	21,188	22,241	22,284	23,296	22,802	24,000	23,420
30%	26,720	26,313	28,121	27,887	29,071	28,636	29,691	29,052
40%	30,979	31,036	32,866	33,187	34,073	34,359	34,556	34,827
50%	35,780	36,321	37,818	38,316	39,014	39,720	40,009	40,215
60%	40,871	42,125	43,405	44,032	44,411	45,705	45,874	46,723
70%	47,449	48,960	49,636	51,295	51,000	52,666	52,372	53,671
80%	55,921	58,317	58,373	60,415	59,643	61,692	61,273	62,872
90%	70,474	73,253	71,347	74,660	73,867	77,437	75,535	78,580

Source: Statistics Austria, EU-SILC 2008-2011. Weighted results.

Figure 1: Absolute and relative deviations of total disposable household incomes at deciles (Register data minus questionnaire data).



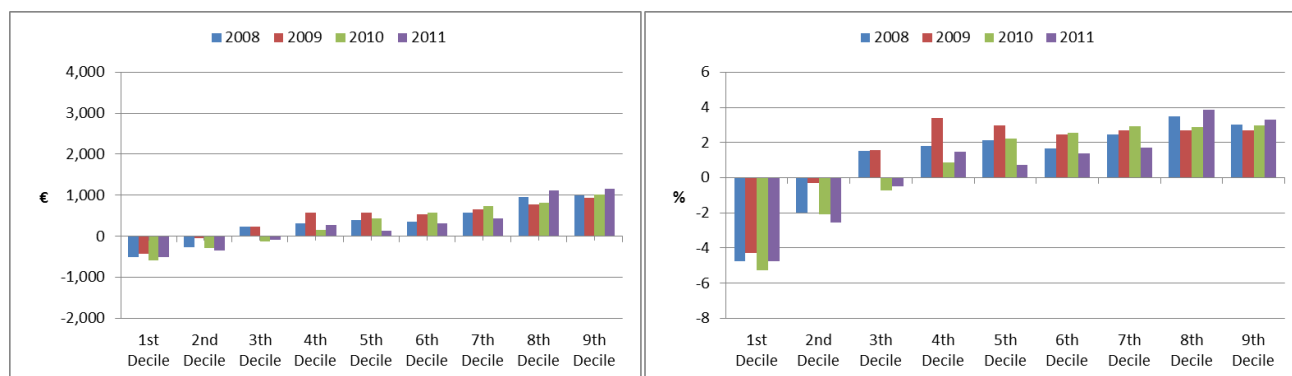
Source: Statistics Austria, EU-SILC 2008-2011. Weighted results.

Table 5: Equivalised disposable incomes 2008 – 2011: questionnaire and register data

	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Count (persons)	8,241,523	8,241,523	8,262,101	8,262,101	8,283,237	8,283,237	8,315,881	8,315,881
Mean	21,379	21,679	22,102	22,750	23,158	23,573	23,642	23,922
Median	19,011	19,413	19,891	20,469	20,618	21,058	21,319	21,463
p80p20	2.05	2.17	2.03	2.09	2.04	2.14	2.04	2.17
p90p10	3.13	3.39	3.06	3.27	3.19	3.46	3.09	3.34
Percentile								
10%	10,653	10,144	11,252	10,817	11,503	10,912	11,898	11,387
20%	13,305	13,037	13,954	13,917	14,388	14,099	14,749	14,392
30%	15,147	15,380	15,880	16,120	16,651	16,533	16,880	16,799
40%	16,996	17,301	17,912	18,498	18,497	18,651	19,016	19,286
50%	19,011	19,413	19,891	20,469	20,618	21,058	21,319	21,463
60%	21,315	21,670	22,143	22,673	22,874	23,443	23,698	24,009
70%	23,901	24,482	24,868	25,524	25,654	26,386	26,466	26,901
80%	27,335	28,284	28,358	29,125	29,388	30,198	30,063	31,187
90%	33,345	34,349	34,454	35,384	36,737	37,759	36,822	37,988

Source: Statistics Austria, EU-SILC 2008-2011. Weighted results.

Figure 2: Absolute and relative deviations of equivalised disposable incomes (Register data minus questionnaire data) at deciles



Source: Statistics Austria, EU-SILC 2008-2011. Weighted results.

Whereas the differences between questionnaire data and register data displayed in Figure 1 and Figure 2 refer to the threshold of a particular decile, Table 6 contains the distribution of differences calculated individually for each household. This can also be interpreted as the observed measurement error for each household if questionnaire data are used. Again, in the lower parts of the distribution we see negative

absolute differences whereas the differences become positive in higher parts of the distribution. If absolute deviations (RD minus questionnaire data) are added up and divided by the questionnaire-based sum of total equivalised disposable income, one gets a proxy for the increase in data completeness due to register data. This indicator amounts to 20% (2008), 19% (2009), 19% (2010), and 15% (2011), respectively.

Table 6: Distribution of differences in equivalised disposable income between register data and questionnaire data

	2008		2009		2010		2011	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Total	158	256	647	328	577	395	596	147
1st Decile	-4,014	-2,298	-4,215	-1,846	-5,194	-2,102	-2,903	-269
2nd Decile	-1,782	-284	-1,255	-118	-1,238	-135	-918	0
3th Decile	-1,157	-17	-656	45	-1,102	-190	-730	0
4th Decile	-669	148	-486	171	-101	324	-219	27
5th Decile	147	451	266	492	-142	359	147	133
6th Decile	-163	266	80	629	46	535	750	419
7th Decile	1,129	1,048	940	889	356	1,032	652	465
8th Decile	917	907	1,158	1,357	1,216	1,394	1,241	887
9th Decile	1,665	1,768	2,470	2,189	2,442	2,202	2,268	1,461
10th Decile	5,513	3,820	8,093	3,783	9,477	4,973	5,667	3,426

Source: Statistics Austria, EU-SILC 2008-2011. Unweighted results.

5.2 DIFFERENCES IN INCOME COMPONENTS

So far, the effects of using register data were discussed for total household income. The aim of this section is to investigate changes in income components in more detail in order to see which parts are most affected by the change of the data source. The focus is on those income types that either amount to a large share of total household income or are particularly important for certain household types (e.g. old age benefits, family/children-related allowances).

Employee Cash or near Cash Income

This type of income refers to the target variable PY010 and has the largest share of total income (around 50% in Austria). For the main part, PY010 is filled with data from the Austrian Wage Tax Dataset. PY010 also includes payments to males in compulsory military or civil service. For the latter, register data about the civil status is available from the dataset on HV Qualifications but, however, does not include the amount of money received. Instead, as their income paid by the government is standardized and adapted each year, it can be derived from other publicly available legal sources.

On the one hand, using register data increases the observed number of income recipients (better coverage). On the other hand, for some income sources register data might lead to undercoverage and missingness. This most likely applies to income not declared for taxation, income from illegal activities and/or income that is not covered in the Wage Tax Dataset because it is not subject to taxation in Austria (e.g.: tip). By using survey instruments, these income sources can potentially be covered. The national questionnaire also includes an additional question that asks for „any other incomes that were not mentioned before“ quoting examples (babysitting, neighborhood help, donations etc.) after all income variables were surveyed. However, it is difficult to estimate the amount and share of non-declared income on total income in SILC in Austria in general. Another issue concerns daily commuters who have their main residence in Austria but are employed in a neighbour country. This mostly occurs in specific border regions. Such incomes are subject to taxation in the foreign country. However, they must be declared in the recipients' annual tax declaration in Austria. Therefore, the Wage Tax Dataset contains some information based on the tax declarations albeit not the actual income which has to be imputed.

Table 7 shows that the use of administrative data for income from employment results in a relatively strong improvement of income coverage. Again, income inequality is higher based on register data and also higher compared to the distribution of total household income and equivalised income. Figure 3 shows that the

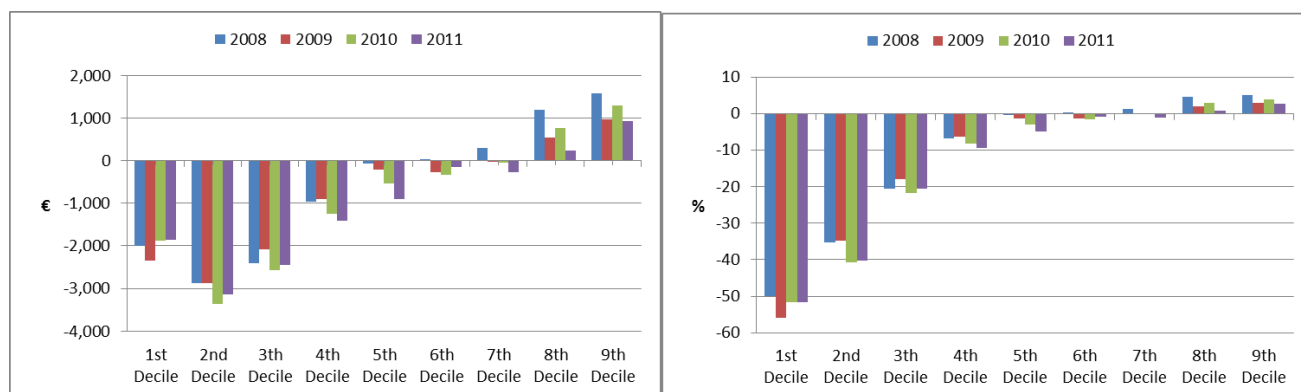
relative deviations between register and questionnaire data for each decile are substantially higher compared to total household income (Figure 1).

Table 7: Employee cash or near cash income (PY010N) 2008 – 2011, questionnaire and register data

	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Count (persons)	3,693,434	3,891,313	3,759,435	4,009,584	3,867,343	4,019,917	3,903,566	4,021,399
Mean	18,275	17,968	18,460	18,352	19,130	18,838	19,185	18,457
Median	16,800	16,735	16,939	16,730	17,500	16,965	18,000	17,104
p80p20	3.13	5.07	3.21	5.03	3.26	5.66	3.53	5.94
p90p10	7.96	16.69	7.83	18.23	9.23	19.87	9.44	20.06
Percentile								
10%	4,002	2,004	4,200	1,856	3,640	1,757	3,600	1,740
20%	8,120	5,250	8,233	5,360	8,238	4,881	7,800	4,668
30%	11,760	9,348	11,628	9,536	11,830	9,265	11,900	9,443
40%	14,400	13,437	14,400	13,500	14,926	13,686	15,176	13,762
50%	16,800	16,735	16,939	16,730	17,500	16,965	18,000	17,104
60%	19,326	19,358	19,739	19,465	20,255	19,929	20,400	20,246
70%	21,938	22,242	22,649	22,629	23,127	23,084	23,600	23,335
80%	25,440	26,636	26,444	26,980	26,855	27,621	27,499	27,725
90%	31,856	33,445	32,871	33,838	33,600	34,904	33,976	34,897

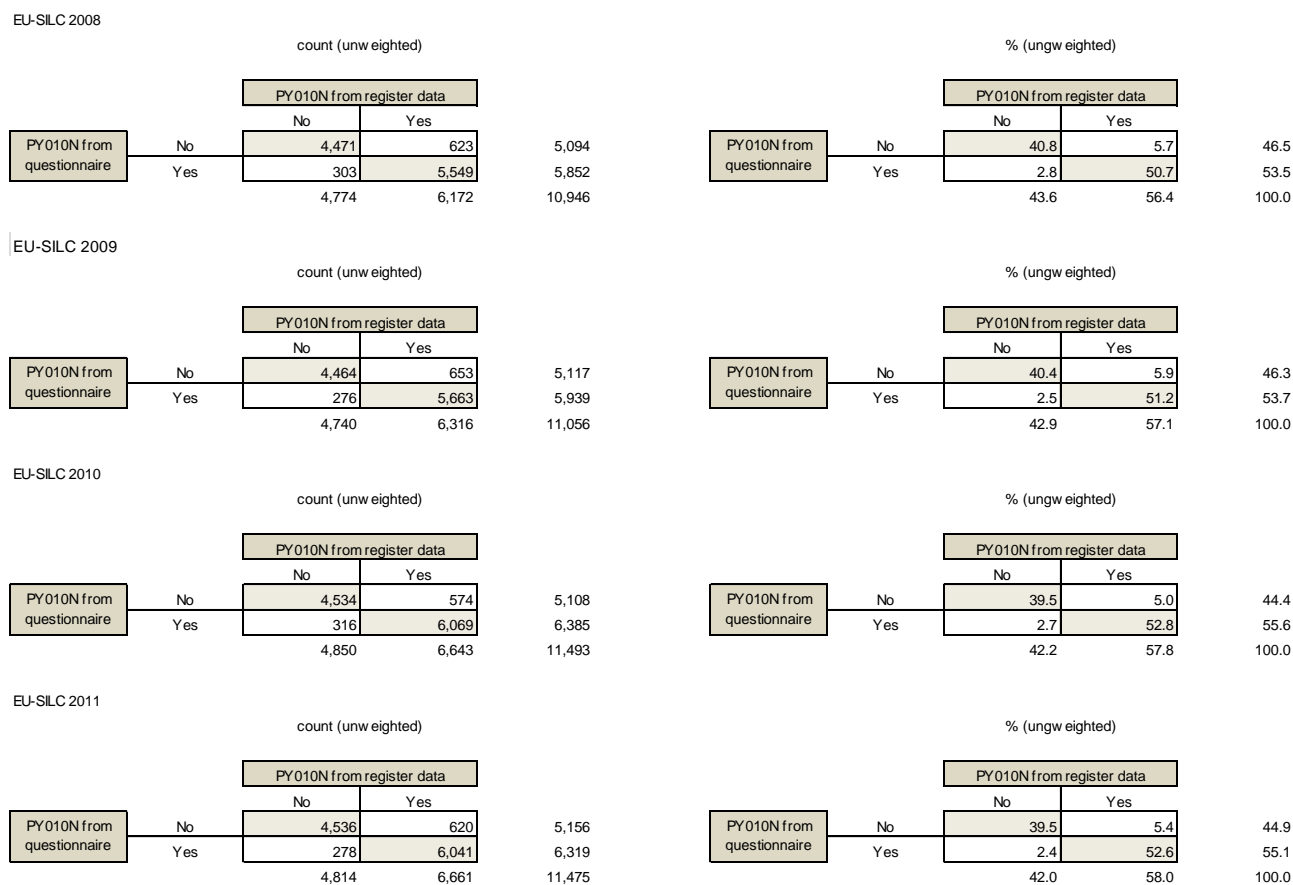
Source: Statistics Austria, EU-SILC 2008-2011. Weighted results.

Figure 3: Absolute and relative deviations of employee incomes (Register data minus questionnaire data) at deciles



Source: Statistics Austria, EU-SILC 2008-2011. Weighted results.

Figure 4: Employee cash or near cash income (PY010N) 2008 – 2011, overlap between data sources



Source: Statistics Austria, EU-SILC 2008-2011. Unweighted results.

Figure 4 depicts the overlap between data sources for PY010N. Although register data are used, there remains some undercoverage. It can be shown that between 276 (EU-SILC 2009) and 316 (2010) persons did not receive any employee income according to register data although they reported such incomes in the questionnaire. Vice versa, between 574 (EU-SILC 2010) and 653 (EU-SILC 2009) individuals had records in the administrative data but did not report any employee income when they were asked in the survey.

In Table 8 the focus is on persons that had employee income according to both questionnaire data and register data. It shows the distribution of people based on whether questionnaire data or register data was used. Thus, some of the observations (all those outside the diagonal) end up with different positions in the income distribution depending on which data source is used. It can be seen that the first two deciles based on register data have a smaller size than other groups. This suggests that individuals whose income is *only* recorded in register data generally receive lower incomes compared to the whole population.

Table 8: Data source and overlap of employee income (PY010N) 2008-2011 per decile (unweighted counts)

Percentiles		Percentile PY010N 2008 with RD										Total
		10	20	30	40	50	60	70	80	90	100	
Percentile PY010N 2008 without RD	10	195	213	47	9	4	6	6	9	1	3	493
	20	38	148	255	44	17	10	2	4	5	4	527
	30	7	39	163	227	56	14	11	3	2	12	534
	40	6	18	37	172	197	76	33	15	3	8	565
	50	4	20	23	53	175	168	81	21	14	10	569
	60	0	8	9	46	89	181	140	69	31	12	585
	70	3	6	10	27	35	81	204	153	42	18	579
	80	1	2	7	12	19	43	90	198	139	40	551
	90	4	6	11	13	5	14	32	109	277	112	583
	100	3	5	8	3	5	10	13	30	96	390	563
Total		261	465	570	606	602	603	612	611	610	609	5,549
Percentiles		Percentiles PY010N 2009 with RD										Total
		10	20	30	40	50	60	70	80	90	100	
Percentiles PY010N 2009 without RD	10	198	216	44	20	7	4	5	4	5	1	504
	20	28	155	265	50	12	4	9	6	3	9	541
	30	11	50	172	233	48	18	8	4	5	6	555
	40	7	16	46	194	202	63	21	14	7	4	574
	50	5	18	24	58	204	172	67	20	11	3	582
	60	4	11	13	32	75	200	152	62	23	8	580
	70	4	7	16	10	35	88	218	129	53	22	582
	80	3	2	6	9	18	37	97	256	127	27	582
	90	0	1	2	2	11	18	33	99	288	131	585
	100	2	4	5	7	5	11	10	24	99	411	578
Total		262	480	593	615	617	615	620	618	621	622	5,663
Percentiles		Percentiles PY010N 2010 with RD										Total
		10	20	30	40	50	60	70	80	90	100	
Percentiles PY010N 2010 without RD	10	241	232	18	12	4	1	0	2	3	2	515
	20	36	184	279	36	5	10	5	5	0	5	565
	30	13	33	217	254	56	18	6	8	6	5	616
	40	10	25	55	239	231	65	15	6	6	7	659
	50	5	14	19	48	189	172	69	9	14	10	549
	60	2	8	13	26	93	234	163	60	19	12	630
	70	4	11	10	13	34	95	270	177	53	17	684
	80	2	3	8	10	19	39	79	264	146	37	607
	90	3	2	8	10	11	21	31	98	323	124	631
	100	5	8	5	3	8	7	20	28	90	439	613
Total		321	520	632	651	650	662	658	657	660	658	6,069
Percentiles		Percentiles PY010N 2011 with RD										Total
		10	20	30	40	50	60	70	80	90	100	
Percentiles PY010N 2011 without RD	10	227	237	32	10	3	1	1	2	3	1	517
	20	42	168	301	31	6	6	5	5	1	5	570
	30	14	36	197	269	55	10	4	1	4	6	596
	40	4	18	41	214	242	74	24	10	2	10	639
	50	4	8	20	66	213	196	74	15	12	6	614
	60	3	9	16	29	73	215	184	73	21	11	634
	70	1	5	6	19	24	80	238	175	49	16	613
	80	2	5	4	11	18	43	93	285	165	42	668
	90	1	6	5	4	9	19	28	76	298	129	575
	100	3	2	6	5	8	15	12	22	107	435	615
Total		301	494	628	658	651	659	663	664	662	661	6,041

Source: Statistics Austria, EU-SILC 2008-2011. Unweighted counts.

Between 37.9% and 42.8% in EU-SILC 2008 to 2011 fall into the same decile group, independent of the data source used. For 44.8% to 46.6% there occurs a deviation by 1 decile (Table 9). Although employee incomes based on register data on average are lower than employee incomes based on questionnaire data, 40% fall into a higher decile when register data are considered. This only holds for 17% to 20% when questionnaire data are used to determine the position in the income distribution (not displayed). In sum, the impact of higher incomes because of additional income components, which were not measured in the questionnaire (e.g. because they were forgotten etc.) is lower than the impact of the higher number of low incomes that are recorded in the register data. As a consequence, this leads to the observed more unequal overall distribution of incomes.

Table 9: Overlap of deciles between register data and questionnaire data

	2008	2009	2010	2011
%-share in the same decile	37.9	40.5	42.8	41.2
%-share deviation by 1 decile	45.3	45.0	44.8	46.6
%-share deviation by 2 deciles	12.4	11.1	10.0	10.4

Source: Statistics Austria, EU-SILC 2008-2011.

Table 10 contains the distribution of differences for PY010N between register data and questionnaire data calculated for each household. Systematic differences between both data sources can be seen. At the bottom of the distribution the difference is negative, i.e. incomes coming from questionnaire data are higher than incomes coming from register data. Among the top deciles the opposite is true. Here, incomes filled from register data are higher. Again, this supports the thesis of the central tendency for reported incomes (section 4.2): higher income-groups report lower amounts and lower income groups report higher incomes.

Table 10: Distribution of differences in employee income (PY010N), register data minus questionnaire data, (unweighted)

	2008		2009		2010		2011	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Total	643.3	392.3	1,249.5	377.2	1,184.4	461.3	627.7	403.9
1st Decile	-2,878.4	-422.5	-3,257.6	-403.4	-3,488.3	-383.4	-2,831.0	-490.4
2nd Decile	-2,829.4	-362.5	-2,902.0	-441.7	-2,971.1	-449.6	-2,619.9	-290.0
3th Decile	-2,279.1	-415.7	-1,586.4	-261.4	-2,074.2	-330.5	-1,614.4	-87.0
4th Decile	-1,652.6	-426.1	-531.6	-45.4	-914.9	-59.9	-1,040.1	-58.4
5th Decile	-247.3	192.4	-413.4	123.4	-900.2	160.5	-498.9	109.9
6th Decile	-190.2	648.9	-28.8	492.3	267.6	602.2	-478.9	441.0
7th Decile	643.7	811.1	480.4	803.9	396.5	912.4	465.7	745.3
8th Decile	575.6	1,238.8	1,341.3	1,046.4	1,257.7	1,504.7	1,301.9	1,222.1
9th Decile	1,937.3	2,007.2	2,455.7	2,226.3	2,860.3	2,622.1	1,805.5	1,923.0
10th Decile	10,301.0	7,354.6	13,200.8	6,615.5	13,969.8	6,931.6	8,938.9	6,329.9

Source: Statistics Austria, EU-SILC 2008-2011. Unweighted results.

The changes yet described focused on the unconditional distribution of income variables. However, structural features of the administrative data can better be made transparent through more detailed analysis. In what follows it is also investigated how changing the data source affects subgroups in the sample. For instance, there could be variations in the differences between register data and questionnaire data if the employment status is considered. Table 11 and Table 12 show whether the overall differences between register and questionnaire data found so far show a similar pattern in selected subgroups. It is evident, that using register data does not affect all subgroups in the same way. On the one hand, the number of persons with a filled value for employee incomes increases if register data is used. The difference is particularly distinctive for households with children. On the other hand, valid data entries are lower for persons that are not employed throughout the year if register data is used.

Table 11: Comparison of employee income (PY010N) 2008 and 2009, by household type and employment status

	2008				2009			
	count (in 1,000's)		Median		count (in 1,000's)		Median	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Households without pension recipients								
Total	3,574	3,756	16,802	17,034	3,642	3,864	17,280	17,156
One person household - males	308	315	20,400	20,420	303	317	21,300	20,602
One person household - females	234	236	18,460	19,385	237	245	18,231	18,786
>1 adult without dependent children	1,334	1,365	16,800	17,853	1,281	1,296	17,530	18,114
Household with dependent children	1,810	1,966	15,600	14,646	1,896	2,080	15,720	14,679
Single parent household, one or more dependent children	112	118	13,200	11,642	128	152	12,900	10,876
>1 adult, one dependent child	767	799	16,520	15,302	844	906	16,069	15,199
>1 adult, two dependent children	610	684	16,944	15,702	622	705	17,683	16,399
>1 adult, three or more dependent children	209	240	16,200	12,551	226	243	15,167	13,517
Employment Status								
employed throughout the year	2,786	2,824	19,600	19,937	2,827	2,840	19,739	20,230
>= 6 months full-time	2,262	2,308	21,083	21,606	2,243	2,262	21,517	22,195
>= 6 months part-time	524	517	11,256	11,409	583	578	11,284	12,228
not employed throughout the year	453	397	8,900	8,429	447	413	11,200	11,000
>= 6 months full-time	249	213	12,950	12,449	286	264	13,758	13,725
>= 6 months part-time	78	66	5,648	6,415	60	55	7,371	7,359
>= 6 months household chores	56	53	3,400	4,560	19	19	3,000	3,357
>= 6 months other activities	71	65	3,970	3,331	82	74	4,400	4,903

Source: Statistics Austria, EU-SILC 2008-2009. Weighted results.

Table 12: Comparison of employee income (PY010N) 2010 and 2011, by household type and employment status

	2010				2011			
	count (in 1,000's)		Median		count (in 1,000's)		Median	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Households without pension recipients								
Total	3,736	3,839	18,000	17,658	3,781	3,866	18,200	17,534
One person household - males	323	332	21,169	20,857	312	315	22,409	21,655
One person household - females	250	249	19,665	19,846	267	262	19,546	19,933
>1 adult without dependent children	1,373	1,364	18,081	18,453	1,400	1,383	18,400	18,860
Household with dependent children	1,808	1,913	16,188	15,063	1,840	1,929	16,520	14,618
Single parent household, one or more dependent children	125	136	13,044	10,660	100	111	15,400	13,440
>1 adult, one dependent child	829	881	16,980	16,528	858	892	16,800	15,351
>1 adult, two dependent children	628	662	16,800	15,933	642	667	16,800	14,652
>1 adult, three or more dependent children	207	216	15,400	11,861	202	237	15,400	12,053
Employment Status								
employed throughout the year	2,754	2,728	20,865	21,171	2,776	2,727	21,069	21,384
>= 6 months full-time	2,214	2,193	22,563	23,043	2,206	2,174	22,869	23,365
>= 6 months part-time	540	534	12,600	12,896	570	553	13,400	13,744
not employed throughout the year	574	545	11,169	10,571	564	520	10,100	10,097
>= 6 months full-time	350	328	14,000	14,126	326	296	14,467	14,299
>= 6 months part-time	94	86	8,000	7,528	88	76	7,010	7,022
>= 6 months household chores	32	33	3,600	3,219	31	31	3,500	3,511
>= 6 months other activities	98	98	4,817	4,375	120	118	3,600	3,271

Source: Statistics Austria, EU-SILC 2010-2011. Weighted results.

The preceding analysis revealed that low incomes are better captured through register data. In what follows, the focus is on those incomes where there is no overlap between both data sources. Table 13 compares incomes that are either available only from questionnaire or administrative data by activity status. It can be seen that those incomes which are recorded in only one of the two data sources are much lower than the overall mean (see Table 7). Incomes from administrative data, in turn, are also substantially lower than incomes from questionnaire data. In 2011, the overall median of those incomes that were only recorded in administrative data amounts to about one third of the median of incomes that were solely recorded in questionnaire data although this difference has steadily decreased over the years. Even persons who report

being employed full-time throughout the year have a lower median income than the total population (see Table 7). Two explanations for that are possible. First, it could indicate that these incomes are not included in the Wage Tax dataset. Second, it could be possible that respondents may believe to be employed, but are not due to missing registration by the employer.⁷ Overall, the proportion of economically active individuals for whom income records are only available from register data is lower than in the complementary group. From 2008 to 2011, more than half of working-age people who have received an employee income only recorded in register data are either predominantly unemployed (≥ 6 months) or economically inactive throughout the year. On the contrary, this applies to only one third of those whose employee income can only be filled with questionnaire data.

Table 13: Employee income (PY010N) by data source and employment status

	2008		2009		2010		2011	
	Questionnaire data only	RD only	Questionnaire data only	RD only	Questionnaire data only	RD only	Questionnaire data only	RD only
count (unweighted)								
Total	303	623	276	653	316	574	278	620
employed throughout the year	108	106	137	94	121	86	106	98
>= 6 months full-time	81	88	98	81	95	77	84	90
>= 6 months part-time	27	18	39	13	26	9	22	8
not employed throughout the year	69	23	47	13	53	31	39	24
>= 6 months full-time	24	7	20	6	19	4	13	6
>= 6 months part-time	13	2	5	2	11	1	8	0
>= 6 months household chores	10	5	3	3	4	9	5	7
>= 6 months other activities	22	9	19	2	19	17	13	11
unemployed (≥ 6 months)	14	59	17	85	30	97	15	105
economically inactive throughout the year	48	257	33	278	53	204	54	243
in retirement	17	70	7	96	19	77	19	75
household chores	17	112	9	54	16	63	13	87
educational activities	14	75	15	69	17	52	21	67
economically inactive due to health problems or other reasons	0	0	2	13	1	12	1	14
	2008		2009		2010		2011	
Median, Euro	Questionnaire data only	RD only	Questionnaire data only	RD only	Questionnaire data only	RD only	Questionnaire data only	RD only
Total	8,400	1,530	7,241	1,539	5,560	1,275	4,827	1,241
employed throughout the year	15,400	3,128	11,200	1,971	15,400	1,755	16,010	1,405
>= 6 months full-time	18,204	3,001	15,255	1,521	19,665	1,838	19,384	1,308
>= 6 months part-time	7,200	3,215	4,883	9,148	5,740	820	5,023	4,048
not employed throughout the year	4,821	2,126	9,392	3,077	4,920	1,189	4,400	1,969
>= 6 months full-time	11,433	8,192	14,350	4,855	11,387	254	8,500	7,892
>= 6 months part-time	5,320	8,741	5,600	9,430	5,717	273	4,513	0
>= 6 months domestic chores	1,458	864	2,400	769	960	1,196	3,500	349
>= 6 months other activities	3,000	682	4,700	3,021	3,680	1,222	3,180	729
unemployed (≥ 6 months)	7,575	1,933	5,167	1,344	4,864	1,793	2,870	1,651
economically inactive throughout the year	3,380	1,978	3,100	2,249	2,100	1,573	3,025	1,416
in retirement	9,800	1,738	4,188	1,528	3,000	1,563	5,110	1,301
household	3,000	1,848	2,400	1,942	2,320	1,757	3,600	1,418
educational activities	2,600	2,248	2,400	1,542	2,000	1,315	2,406	1,473
economically inactive due to health problems or other reasons	0	0	4,800	2,439	2,100	2,577	12,000	1,510

Source: Statistics Austria, EU-SILC 2008-2011. Weighted results.

Unemployment Benefits

Income from unemployment benefits in Austria are publicly organised in a social insurance and are paid by the Austrian Public Employment Service (AMS). This is an agency under public law that provides labour-

⁷ Persons who are employed (based on self-assessment), but would count as self-employed if a legal definition would be strictly followed.

market related services. All payments are summarized in the EU-SILC target variable PY090. Although unemployment benefits account for only about 2% of total household income (Table 2), they are of significant economic relevance for some groups of people/households.

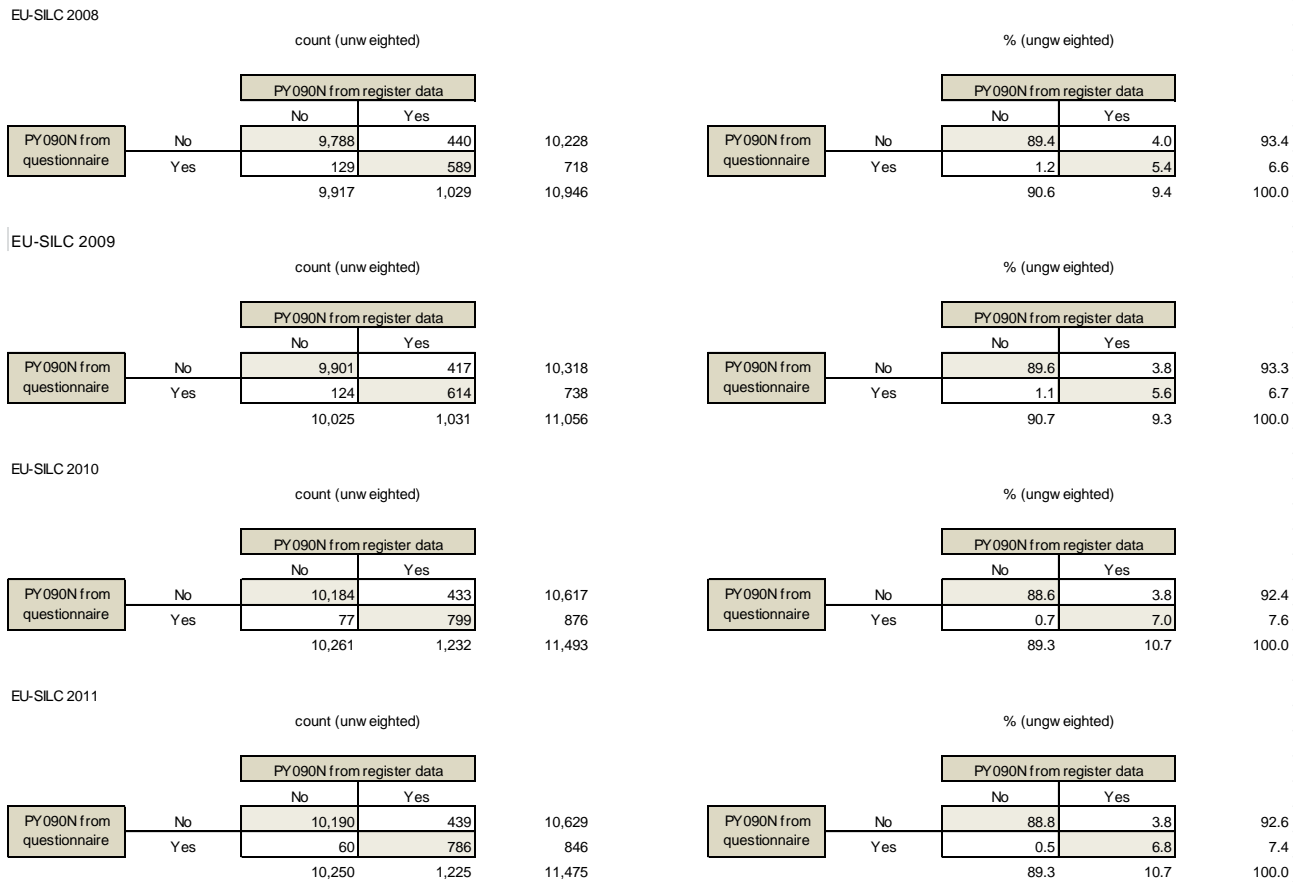
For filling PY090 the Transfers Dataset is used. It contains data on unemployment benefits (daily rates and total for each unemployment spell) and the beginning and ending date of unemployment spells (resulting in the beginning and terminating of benefits). Each record corresponds to one unemployment spell. Thus, aggregation is necessary. PY090 contains around 30 different types of allowances for people out of employment. As unemployment benefits could possibly be paid only for a very short time period they are particularly likely to be under-reported in a survey situation. Comparing questionnaire data and register data reveals a better coverage of unemployment benefits in the latter (Table 14, Figure 5).

Table 14: Unemployment benefits (PY090N) 2008 – 2011): questionnaire and register data

	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Count (persons)	570,264	698,694	555,471	707,624	671,757	876,178	644,314	841,798
Mean	4,300	3,395	4,130	3,385	4,125	3,756	4,625	3,910
Median	3,000	2,303	2,940	2,316	3,000	2,754	3,250	2,744
Percentile								
10%	720	456	780	432	800	604	802	541
20%	1,200	923	1,260	798	1,340	1,058	1,400	1,019
30%	1,640	1,322	1,760	1,202	1,940	1,662	2,000	1,594
40%	2,240	1,801	2,250	1,664	2,400	2,157	2,580	2,166
50%	3,000	2,303	2,940	2,316	3,000	2,754	3,250	2,744
60%	4,200	2,988	3,990	2,920	3,920	3,414	4,380	3,557
70%	5,575	4,097	5,142	3,825	5,000	4,471	6,000	4,771
80%	7,200	5,464	7,080	5,569	6,600	6,071	7,620	6,350
90%	9,000	7,495	8,760	7,968	9,504	8,321	9,800	8,701

Source: Statistics Austria, EU-SILC 2008-2011. Weighted Results.

Figure 5: Unemployment benefits (PY090N) 2008 – 2011, overlap between data sources



Source: Statistics Austria, EU-SILC 2008-2011. Unweighted results.

If only those unemployment benefits are investigated which are available in both data sources, it can be shown that the distribution of benefits is only marginally different between questionnaire and register data (Table 15). Payments that are recorded solely in the Transfer dataset, however, are considerably lower, suggesting that the majority of these payments only last for short time spells. Benefits which exclusively come from questionnaire data are generally higher (in the top percentile even higher than the maximum unemployment benefit).

Table 15: Distribution of unemployment benefits PY090N 2008-2010 by data source

	2008				2009			
	Questionnaire data only	PY090N available in both sources		RD only	Questionnaire data only	PY090N available in both sources		RD only
		Questionnaire	RD			Questionnaire	RD	
Count (persons)	129	589	589	440	124	614	614	417
Mean	6,152	3,853	3,891	2,646	6,359	3,727	3,760	2,697
Median	5,268	2,800	2,892	1,707	4,300	2,745	2,636	1,366
Percentile								
10%	657	690	793	275	900	720	678	258
20%	1,200	1,100	1,267	528	1,500	1,200	1,170	432
30%	1,800	1,600	1,787	895	2,250	1,555	1,618	773
40%	2,880	2,100	2,251	1,225	3,200	2,100	2,218	989
50%	5,268	2,800	2,892	1,707	4,300	2,745	2,636	1,366
60%	6,600	3,740	3,719	2,172	5,400	3,548	3,362	1,924
70%	7,600	4,900	4,577	2,820	6,935	4,500	4,364	2,747
80%	9,369	6,600	6,111	3,831	9,360	6,360	5,923	3,847
90%	13,948	8,400	8,434	6,202	14,640	8,290	8,544	6,909

	2010				2011			
	Questionnaire data only	PY090N available in both sources		RD only	Questionnaire data only	PY090N available in both sources		RD only
		Questionnaire	RD			Questionnaire	RD	
Count (persons)	77	799	799	433	60	786	786	439
Mean	5,013	4,059	4,284	2,704	7,184	4,443	4,467	2,749
Median	3,600	2,940	3,193	1,703	5,910	3,200	3,414	1,634
Percentile								
10%	760	750	824	283	1,150	800	810	292
20%	1,200	1,200	1,435	580	1,950	1,350	1,511	537
30%	2,000	1,800	2,047	893	2,480	1,880	2,111	836
40%	2,700	2,400	2,548	1,249	4,320	2,500	2,713	1,254
50%	3,600	2,940	3,193	1,703	5,910	3,200	3,414	1,634
60%	4,209	3,900	4,120	2,177	7,510	4,080	4,174	2,235
70%	5,160	4,992	5,341	2,884	8,928	5,760	5,516	2,816
80%	7,560	6,640	7,030	3,996	12,783	7,200	6,778	4,204
90%	10,425	9,240	9,157	6,615	17,238	9,600	9,360	7,176

Source: Statistics Austria, EU-SILC 2008-2011. Unweighted results.

Old-age Benefits

Old-age benefits are captured in the EU-SILC target variable PY100. In general, this refers to benefits that should compensate for the loss of income due to age-related retirement from work. PY100 includes pension benefits from public social insurance as well as care allowances and similar age-related social transfers. As old-age benefits comprise a broad range of different benefits multiple register datasets (Chapter 0) have to be used.

Unlike employee incomes, the differences in the distribution of old-age benefits between questionnaire and register data are not very high. Income inequality based on register data is slightly greater than for questionnaire data (Table 16), although the difference is considerably smaller than for employee income (Table 7). The reduction of recipients between 2010 and 2011 based on questionnaire data occurs because survivor' benefits received below the standard retirement age were not included anymore in PY100 starting from 2011 but calculated in the associated target variable.

Table 16: Old-age benefits (PY100N) 2008 – 2011: questionnaire and register data

	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Count (persons)	1,711,659	1,654,527	1,726,821	1,703,951	1,753,052	1,757,354	1,677,305	1,754,813
Mean	15,987	15,807	16,574	16,362	17,027	17,151	16,875	16,308
Median	14,924	15,112	15,395	15,643	15,760	16,247	15,759	15,355
p80p20	2.29	2.60	2.32	2.59	2.35	2.72	3.09	3.43
p90p10	4.21	5.32	4.27	5.38	4.29	5.73	5.87	6.99
Percentile								
10%	6,158	4,847	6,223	4,954	6,369	4,931	4,897	4,108
20%	9,369	8,394	9,542	8,666	9,733	8,763	7,777	6,922
30%	10,899	10,764	11,096	11,344	11,802	11,655	10,416	9,499
40%	13,158	13,227	13,623	13,532	13,947	14,145	13,068	12,298
50%	14,924	15,112	15,395	15,643	15,760	16,247	15,759	15,355
60%	16,822	17,127	17,380	17,715	18,384	18,661	18,270	17,921
70%	18,752	19,122	19,868	19,734	20,441	20,975	20,843	20,607
80%	21,498	21,829	22,104	22,437	22,825	23,817	24,028	23,710
90%	25,952	25,781	26,570	26,644	27,323	28,260	28,767	28,699

Source: Statistics Austria, EU-SILC 2008-2011. Weighted Results.

Family/Children-related allowances

Family/Children-related allowances are summarized in the EU-SILC target variable HY050. It amounts to approximately 5% of the total household income. Similarly to unemployment benefits this is a rather small share, but for households with children these allowances can be very important in order to remain above the poverty threshold. The primary administrative data source is the National Dataset on Family/Children-Related Allowances. Table 17 compares the distribution of family-related benefits for the years 2008-2011. Differences between questionnaire data and register data closely follow the overall trend. Dissimilarities between questionnaire data and administrative data for HY050 are lower than for employee incomes PY010N.

Table 17: Family/Children-related Allowances (HY050) 2008 – 2011: questionnaire and register data

	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Count (households)	1,172,058	1,117,694	1,158,367	1,160,596	1,147,630	1,116,125	1,167,386	1,180,768
Mean	4,935	4,926	5,460	5,199	5,344	5,180	5,491	5,564
Median	4,298	4,298	4,675	4,511	4,735	4,607	4,735	4,735
Percentile								
10%	2,127	1,963	2,363	2,076	2,339	2,166	2,377	2,166
20%	2,182	2,182	2,363	2,313	2,403	2,403	2,403	2,403
30%	2,400	2,400	2,647	2,356	2,721	2,475	2,672	2,686
40%	3,283	3,841	4,420	3,967	3,520	3,944	3,252	4,200
50%	4,298	4,298	4,675	4,511	4,735	4,607	4,735	4,735
60%	4,517	4,517	4,893	4,791	4,971	4,971	4,971	4,971
70%	5,040	5,246	6,520	5,487	5,573	5,467	5,573	6,296
80%	7,223	7,108	8,025	7,433	7,939	7,556	7,977	7,683
90%	9,365	9,365	10,176	9,809	9,872	9,982	10,408	10,836

Source: Statistics Austria, EU-SILC 2008-2011, calculations for households. Weighted results.

5.3 DIFFERENCES IN THE COMPOSITION OF THE TOTAL HOUSEHOLD INCOME

The total disposable household income is the aggregation of all income components. In this section, the structure of the household income and the relative importance of its components are investigated and compared for both questionnaire data and register data. This can be done in two ways. First average shares of major components are tabulated. Second, it is checked how these shares (which themselves vary over households) are distributed among different groups of households.

Average shares of major income components for the whole sample (column labelled 'total') are almost the same over the years 2008 – 2011 and also do not depend on the data source used (Table 18). However, marked differences are observable between households at risk of poverty and households not at risk of poverty. Roughly speaking, for the period 2008 – 2011 the share of employee incomes on total income for households not at risk of poverty is twice as high as for households at risk of poverty. Conversely, social transfers amount to more than two thirds of total income for the latter and to approx. 10% for the former. This relation does not alter very much if one switches from questionnaire data to register data. For 2008 and 2009 the share of employee incomes for households at risk of poverty is slightly higher when register data is used. However, this difference diminished after 2009.

Table 18: Share of income components on household income, by poverty status

	2008						2009					
	Questionnaire			RD			Questionnaire			RD		
	Total	Not AROP	AROP	Total	Not AROP	AROP	Total	Not AROP	AROP	Total	Not AROP	AROP
Employee income	57	59	30	58	60	34	57	58	30	58	60	34
Income from self-employment	8	8	7	9	9	7	9	9	6	9	9	8
Social transfers	11	10	34	10	9	31	11	10	32	11	10	30
Old-age benefits	20	19	26	18	18	23	20	19	28	19	18	23
Private transfers	4	4	3	4	4	4	3	3	3	3	3	4

	2010						2011					
	Questionnaire			RD			Questionnaire			RD		
	Total	Not AROP	AROP	Total	Not AROP	AROP	Total	Not AROP	AROP	Total	Not AROP	AROP
Employee income	57	59	27	57	59	28	56	58	28	55	57	28
Income from self-employment	9	9	6	9	9	7	9	9	7	9	9	9
Social transfers	11	9	36	11	10	37	12	11	42	13	12	40
Old-age benefits	19	19	27	19	19	23	18	18	20	18	18	19
Private transfers	4	4	4	4	4	5	4	4	4	4	4	5

Source: Statistics Austria, EU-SILC 2008-2011. Weighted results. AROP = at risk of poverty.

In a second step the focus is on the distribution of the %-share on total income among households. In order to have a readable table, results are presented for employee income only. As it can be seen in Table 19, the Median of the %-share of employee income on total household income among households at risk of poverty is higher if register data are used. Particularly for the years 2008 and 2009 this difference is rather big. Furthermore, the %-share of employee income recipients among all households at risk of poverty is higher if register data are used. These two observations, in part, are a consequence of the better coverage of lower employee incomes in register data. The AROP-group also has a lower absolute median for employee income. In sum, on the one hand we see a higher share of households that receive employee income based on register data among those that are at risk of poverty as compared to questionnaire data. On the other

hand, the absolute value (median) of this income decreases. By using register data one gets a slightly different household income structure for households at risk of poverty.

Table 19: Share of employee income on total household income by poverty status

	2008						2009					
	Questionnaire			RD			Questionnaire			RD		
	Total	Not AROP	AROP	Total	Not AROP	AROP	Total	Not AROP	AROP	Total	Not AROP	AROP
Median of %-share (including those where %-share=0)	63	68	2	64	69	15	64	68	6	64	69	18
median of %-share of income recipients (excluding those where %-share=0)	78	79	63	78	80	59	78	79	60	78	79	55
%-share of recipients of employee income	76	79	50	77	81	58	76	79	52	78	81	60
median of employee income of recipients	28,200	30,091	11,282	28,214	30,992	9,754	29,680	31,488	9,312	29,899	32,535	8,599

	2010						2011					
	Questionnaire			RD			Questionnaire			RD		
	Total	Not AROP	AROP	Total	Not AROP	AROP	Total	Not AROP	AROP	Total	Not AROP	AROP
Median of %-share (including those where %-share=0)	63	68	6	61	68	9	62	68	2	58	65	9
median of %-share of income recipients (excluding those where %-share=0)	77	79	49	76	78	44	77	79	51	74	77	42
%-share of recipients of employee income	76	80	52	78	81	58	76	80	50	77	80	60
median of employee income of recipients	29,980	32,262	8,983	30,082	32,773	7,443	31,072	33,154	10,339	29,616	33,180	8,331

Source: Statistics Austria, EU-SILC 2008-2011. Weighted results. AROP = at risk of poverty.

5.4 DIFFERENCES IN EU 2020 INDICATORS

The main EU 2020 indicators in the field of social inclusion are the number of people at risk of poverty or social exclusion and its three components: persons at risk of poverty after social transfers, persons severely materially deprived, persons aged 0-59 living in households with very low work intensity. The latest figures for 2013 show that 1.572 million people, or 18.8% of the total population, in Austria were at risk of poverty or social exclusion. This means that these people were in at least one of the following three conditions: at-risk-of-poverty after social transfers (income poverty), severely materially deprived or living in households with very low work intensity. Looking at each of the three elements contributing to being at risk of poverty or social exclusion, 14.4% of the national population in 2013 were at risk of poverty after social transfers, meaning that their disposable income was below the national at-risk-of-poverty threshold which is set at 60% of the national median equivalised disposable income (after social transfers). In 2013, this threshold amounts to 1,104 Euro per month for one person households, plus 331 Euro for each child below the age of 14, plus 552 Euro for each adult in the household. 4.2% of the population were severely materially deprived, meaning that they had living conditions constrained by a lack of resources (at least 4 out of 9 indicators) such as not being able to afford to pay their bills, keep their home adequately warm, or take a one week holiday away from home. 7.7% of the population aged 0-59 in the Austria lived in households where the adults in working age worked less than 20% of their total work potential during the past year. These three conditions can also occur in combination. In 2013, around one fourth of the people at risk of poverty or social exclusion or 4.6% of the total population had at least 2 out of 3 conditions. This corresponds to 385.000 persons. 97.000 faced an overlap of all three conditions (1% of total population). The magnitude of the EU 2020 indicators for Austria generally lies below the EU 28 average. In the EU, 122.6 million people, or 24.5% of the population, were at risk of poverty or social exclusion in 2013. 17% of the population in the EU28 was at risk of poverty and 10% were severely materially deprived. 11% lived in households with very low work intensity.

Changes in the distribution of the household income and equivalised income as described in the previous sections eventually also affect income-based indicators like the Gini coefficient or the at-risk-of-poverty rate. Due to a change in the calculation of weights, also non-income based indicators like (severe) material deprivation can be affected. In this chapter it is checked how the usage of register data impacts on these EU 2020 indicators for the time period 2008 to 2011.

A more unequal income distribution was repeatedly described as a result of the integration of administrative data in EU-SILC in the previous sections. This is reflected in the S80/S20 measure (sum of the equivalent

income of the highest income quintile divided by sum of the equivalent income of the lowest income quintile) and the Gini coefficient (0 %: total equipartition, 100%: total concentration of income to an individual case) which exhibit a higher value if register data are used (Table 20).

Table 20: S80/S20 and Gini indices 2008-2011

	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
S80/S20	3.7	4.2	3.7	4.2	3.7	4.3	3.8	4.1
Gini	26.1	27.7	25.6	27.5	26.1	28.3	26.3	27.4

Source: Statistics Austria, EU-SILC 2008-2011. Based on weighted data.

As the distribution of the equivalised household income has become more unequal but the median basically remains unchanged after the usage of register data (Table 5) more people fall below the poverty threshold. Consequently, we observe a higher absolute and relative number of persons at risk of poverty in all four years 2008 – 2011 (Table 21). The difference, measured in percentage points, is highest in 2008 (2.8 percentage points) and lowest in 2011 (0.7 percentage points). Table 21 also shows that there are only minor differences in the poverty threshold between questionnaire data and register data. The deviation of the absolute threshold, measured as percentage of the poverty threshold based on questionnaire data, lies between 0.7% (2011) and 2.9% (2009). Moreover, the measure for the poverty gap rises (with the exception of the year 2011).

Table 21: Persons at-risk-of-poverty after social transfers 2008-2011

	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
at-risk-of-poverty after social transfers, %	12.4	15.2	11.9	14.5	12.1	14.7	12.6	14.5
at-risk-of-poverty after social transfers (persons)	1,018,472	1,252,237	983,656	1,201,025	1,003,935	1,213,915	1,051,383	1,206,894
60% risk-of-poverty threshold, Euro	11,406	11,648	11,935	12,281	12,371	12,635	12,791	12,878
poverty gap, %	15.3	19.9	16.9	19.2	17.2	21.8	19.0	19.1

Source: Statistics Austria, EU-SILC 2008-2011. Based on weighted data.

The percentage of people at risk of poverty or social exclusion increased by roughly 2% in each year after the back calculation with register data as compared to questionnaire data (Table 22). As the two other non-income based indicators (severe material deprivation, low work intensity) are only affected through the newly calibrated weights there occur only minor differences to the questionnaire-based results.

Table 22: Europa 2020 indicators 2008 - 2011

frequencies (in 1,000s, weighted)	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
at risk of poverty or social exclusion	1,535	1,699	1,425	1,577	1,373	1,566	1,407	1,593
at risk of poverty after social transfers	1,018	1,252	984	1,201	1,004	1,214	1,051	1,207
living in households with very low work intensity	506	475	487	452	497	497	519	546
severely materially deprived	524	485	395	376	356	353	325	333

in %	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
at risk of poverty or social exclusion	16.6	18.9	16.6	18.9	16.9	19.2	16.9	19.2
at risk of poverty after social transfers	12.1	14.7	12.1	14.7	12.6	14.5	12.6	14.5
living in households with very low work intensity								
as % of the total population	6.0	6.0	6.0	6.0	6.2	6.6	6.2	6.6
in % of people with age < 60 ¹	7.7	7.8	7.7	7.8	8.0	8.5	8.0	8.5
severely materially deprived	4.3	4.3	4.3	4.3	3.9	4.0	3.9	4.0

Source: Statistics Austria, EU-SILC 2008-2011. Based on weighted data.

Table 23 shows the at-risk-of-poverty rate differentiated by date source and household type. There is no single obvious pattern that could explain the observed variations between questionnaire data and register data. In general, the increase in the change in the share of persons at risk of poverty in each group corresponds to the overall increase for the total population if register data are used. This change is slightly

stronger for persons in households with 3 or more children and for females without pension living alone compared to households with pension recipients and persons in households with less than three children.

Table 23: At-risk-of-poverty rate by household type 2008 - 2011

	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Total	12	15	12	15	12	15	13	15
Households with pension recipients								
Total	16	17	16	16	15	15	13	13
One person household - males	15	14	11	12	13	12	13	13
One person household - females	24	26	28	26	26	27	26	26
>1 adult	13	14	12	12	11	11	10	10
Households without pension recipients								
Total	12	15	11	14	11	15	12	15
One person household - males	16	21	17	22	19	23	23	26
One person household - females	20	25	17	25	24	30	26	28
>1 adult without dependent children	6	10	7	10	6	9	7	9
Household with dependent children	13	15	11	14	12	15	12	15
Single parent household, one or more dependent children	29	31	29	27	28	31	24	26
>1 adult, one dependent child	9	9	7	9	7	7	5	8
>1 adult, two dependent children	10	12	7	9	11	13	10	14
>1 adult, three or more dependent children	20	27	20	32	18	29	26	25

Source: Statistics Austria, EU-SILC 2008-2011. Weighted results.

There also occur discrepancies between questionnaire and register data if the poverty risk is differentiated by education and citizenship. The poverty rate for people with Non-Austrian citizenship is markedly higher if register data are used. This particularly applies to persons with EU/EFTA citizenship for the years 2008 and 2010. Similar differences can be found if the ISCED level is considered although these dissimilarities (measured in percentage points) become smaller over time for most education groups.

Table 24: At-risk-of-poverty rate by citizenship and highest education level (PE040) 2008 -2011

%	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Citizenship								
Austria	11	12	10	12	10	12	11	12
born outside Austria	21	30	25	28	26	29	23	21
Non-Austrian								
total	26	36	24	34	25	36	29	34
EU/EFTA	16	31	20	32	15	33	23	30
Other countries	30	39	25	35	31	37	32	36
Highest ISCED level attained								
pre-primary or primary	32	43	24	36	42	45	43	38
lower secondary	22	23	21	22	20	21	20	21
(upper) secondary	10	12	10	12	9	11	10	12
post-secondary non tertiary education	6	10	7	10	6	8	7	8
tertiary	6	10	5	8	7	10	6	9

Source: Statistics Austria, EU-SILC 2008-2011. Weighted results.

Working poor

As was shown in previous sections, using register data changes the observed distribution of employee income (5.2) and the composition of household income (5.3). Consequently, we also observe an effect on the number and share of persons that are working poor. As can be seen in Table 25 based on register data there are more people working poor. The difference amounts to around two percentage points in the years 2008 - 2011. Table 26 displays the composition of the working poor differentiated by work status. Dissimilarities between the two data sources are rather small and mainly occur for people working part time.

Table 25: Working poor 2008-2011

	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Persons aged 18 to 64*	227,352	307,574	217,507	300,860	181,003	269,593	197,745	274,409
% of people aged 18 to 64*	6.4	8.5	6	8.2	5	7.5	5.4	7.6

Source: Statistics Austria, EU-SILC 2008-2011. * Persons that were working for at least 6 months in the preceding year. Weighted results.

Table 26: Working poor by employment status 2008 - 2011

% shares on total no. of working poor	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Year-round employment	81	80	75	77	74	74	73	76
of which at least 6 months								
full time	58	56	54	57	52	56	48	56
part time	23	23	21	20	22	18	25	21
No year-round employment	19	20	25	23	26	26	27	24
of which at least 6 months								
full time	14	15	20	18	16	19	20	18
part time	4	5	4	3	9	6	6	5
other	0.5	0.0	1	1	1	0.5	1	1

count (weighted)	2008		2009		2010		2011	
	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD	Questionnaire	RD
Year-round employment	179,106	237,254	160,410	226,027	127,076	190,808	138,683	200,026
of which at least 6 months								
full time	127,504	167,862	115,151	168,062	89,262	143,500	91,963	145,623
part time	51,602	69,392	45,259	57,965	37,814	47,308	46,720	54,403
No year-round employment	40,664	60,209	53,336	68,358	45,640	67,063	51,526	62,097
of which at least 6 months								
full time	30,782	45,419	43,290	54,038	28,340	49,773	38,210	48,326
part time	8,866	14,790	8,461	10,207	16,028	16,050	11,439	11,988
other	1,016.4	0	1,585	4,113	1,272	1,240.2	1,877	1,783

Source: Statistics Austria, EU-SILC 2008-2011. Weighted results.

6. CONCLUSIONS

This report aims at demonstrating the effects of the substitution of questionnaire data by administrative data for selected income variables and EU 2020 indicators for the time period 2008-2011. Since the beginning of EU-SILC in Austria in 2003, questionnaires were used as survey instruments for collecting income data from private households. A new national data regulation in Austria in 2010, however, allowed using administrative data for income measurement in EU-SILC. After several checks and explorations of methodical limits and possibilities, administrative data have been used on a regular basis since SILC 2012 to construct income target variables and survey weights. As a consequence, this created a break in time series between 2011 and 2012. In order to shift back this break to the base year of the EU 2020 targets (2008), it was decided to do a comprehensive back-calculation of EU-SILC in Austria using administrative data for 2008, 2009, 2010 and 2011. This process was completed by mid-October 2014. Final data have been sent to Eurostat in order to revise Austrian data from 2008 to 2011.

The comparison of results based on questionnaire data with results based on register data leads to three main conclusions: first, the distribution of household incomes in Austria becomes more unequal if income from administrative data is used. This is mainly due to a better coverage of incomes at the bottom and top of the distribution. Second, average shares of major income components for the whole sample (column labelled 'total') remain almost the same over the years 2008 – 2011 and also do not depend on the data sources used. Third, as the median and poverty threshold barely change this leads to a higher at-risk-of-poverty rate (and a higher at-risk-of-poverty-or-social exclusion rate) for Austria for the years 2008 – 2011. The non-income based EU 2020 indicators, severe material deprivation and low work intensity are also affected through the newly calibrated weights but show only minor differences compared to the questionnaire-based results.

Concerning the use of Austrian EU-SILC data by the scientific community a clear communication about the implications of using registers instead of questionnaires for filling income variables and calculating weights is important.

7. REFERENCES

Glaser, Th./Till, M. (2010), Gewichtungsverfahren zur Hochrechnung von EU-SILC-Querschnitt-ergebnissen. In: Statistische Nachrichten 7/2010. Statistik Austria. Wien. pp 566-577.

Statistik Austria. (2013), Methodenbericht EU-SILC 2012. Wien.

The Canberra Group (2001). Expert Group on Household Income Statistics: Final Report and Recommendations. Ottawa.

The Canberra Group (2011). Handbook on Household Income Statistics. Second Edition.