Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
**Federal Statistical Office FSO**


# Construction of Full Time Equivalent for the Swiss Business Frame

Monique Graf and Jann Potterat

Swiss Federal Statistical Office, Statistical Methods Unit

European Establishment Statistics Workshop
EESW11 12-14 September 2011, Neuchâtel, Switzerland

# Outline

# Introduction


Swiss business census will be replaced by the use of registers and complementary surveys.

Main sources:

1. Business register (BR)
   - records the new businesses
   - updates the economic activity
2. Social security register (SR)
   - provides information about gender, no of employees and wages at the level of the enterprise
3. Quarterly Survey of Employment (JobStat)
   - provides full time equivalents (FTE) on a quarterly basis for approx. 36 000 enterprises

# Record linkage of BR and SR


The record linkage of the two registers is ongoing and should be completed in 2013 for the year 2011.

- The information should be available in the future within a two years time lag.
- Full time equivalents (FTE) are missing
- Very large companies are asked directly for FTE.

The FTE predictive model will be applied to the remaining companies.

# Information at disposal


1. Information from the register (=BR+SR)
   - Characteristics of the wage distribution by gender and month
   - Total employment by gender and month
   - economic activity
   - region
2. Information from JobStat
   - FTE by gender on 1 day per quarter
   - Total employment by gender on 1 day per quarter
   - economic activity
   - region

# Method principle


Construct FTE by model imputation:

- Set up a predictive model using the information on the enterprises participating to JobStat
- Use it to impute FTE for the other enterprises in the register.

# Monthly wage distribution


Information from the register:

- Company level: company's economic activity and region;
- Employee level: working months, annual salary and gender.

Using this information, we can estimate an average monthly wage for each employee by dividing the annual wage by the no of working months.

$\Rightarrow$ Company's wage distribution by gender for a given month.

# Model variables by quarter and by gender


Instead of predicting the FTE directly, we predict the mean occupation level (MOL) per gender in the company.

$$\text{MOL} = \frac{\text{FTE}}{\text{TOT}}$$

where TOT is the total employment by gender in JobStat.
Explicative variables from the company's wage distribution:

- mean of log(wage) by gender : MW
- standard deviation of log(wage) by gender : SdW
- skewness coefficient of log(wage) by gender : SkW

Other explicative variables : NACE2 and NUTS2.

# Estimation (mass imputation)

Estimation is done using a mixed model via the SAS procedure TRANSREG:

$$\begin{aligned} y = \log(\text{MOL}) &= \text{OPSCORE(NUTS2)} + \text{OSPCORE(NACE2)} \\ &+ \text{IDENTITY(MW SdW SkW)} \end{aligned}$$

Then this predictive equation is applied to all non sampled companies in the register.

$$\hat{y} = \widehat{\log(\text{MOL})}$$

$$\Rightarrow \qquad \widehat{\text{FTE}} = \exp(\hat{y})\text{TOT'},$$

where TOT' is the total employment by gender in the register.

# Validation


- The Swiss Earnings Structure Survey (SESS) is a biennial business survey that provides data on the occupancy rates and wages.
- The SESS makes it possible to test the feasibility of a FTE model with all the information on the same source.
- The SESS wage and type of workers definition is not totally the same as in the SR.

The performance of the model can nevertheless be tested on the SESS.

# Overall quality


Two models have been considered:

- Model 1: the estimation is done separately in 4 groups (sectors x 2 NACE groups with high or low median wage);
- Model 2: one model for the whole economy.

We show the relative difference between

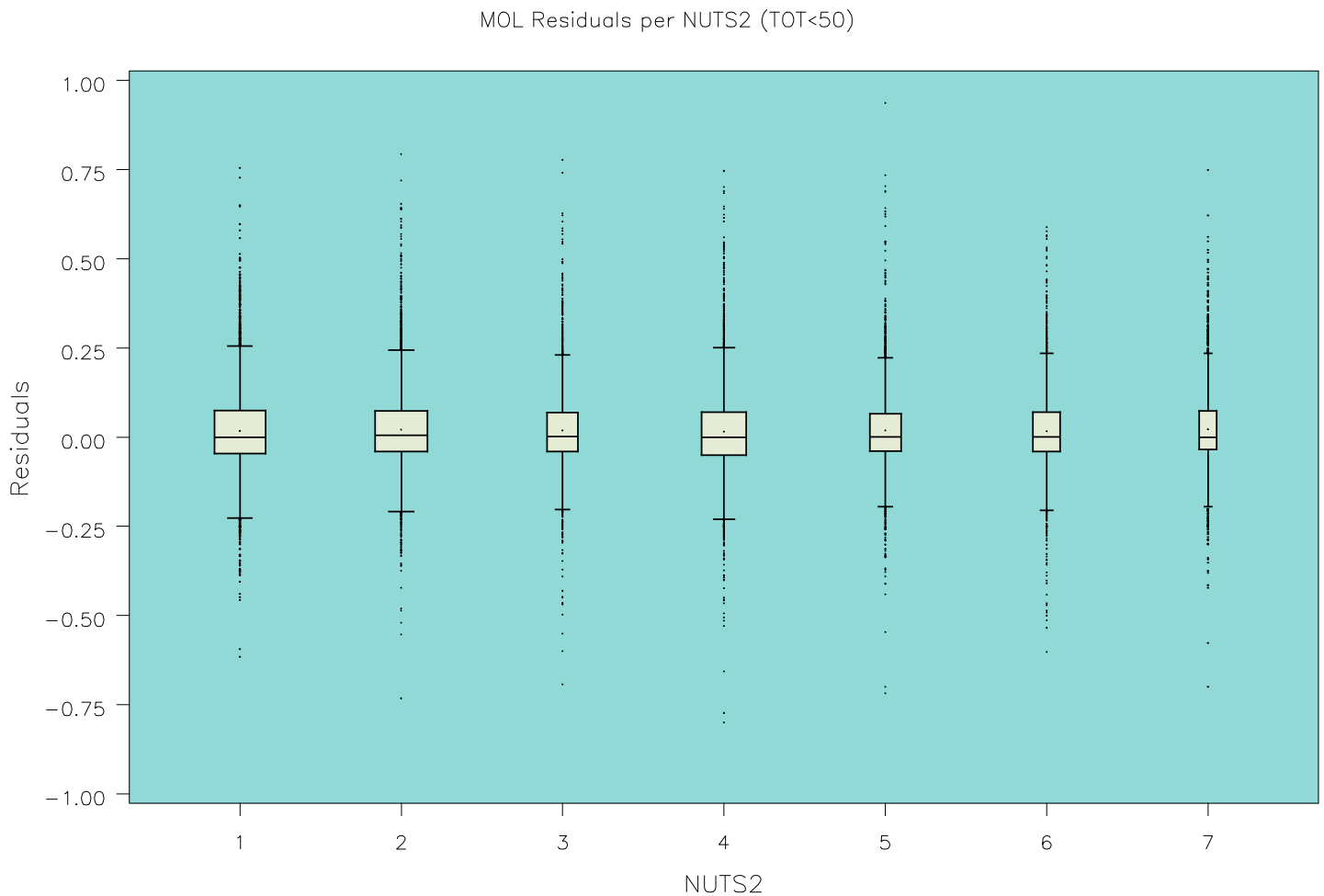
- the total FTE by NACE2 as observed in the SESS (benchmark) with
- the total given by the model.

# Overall quality

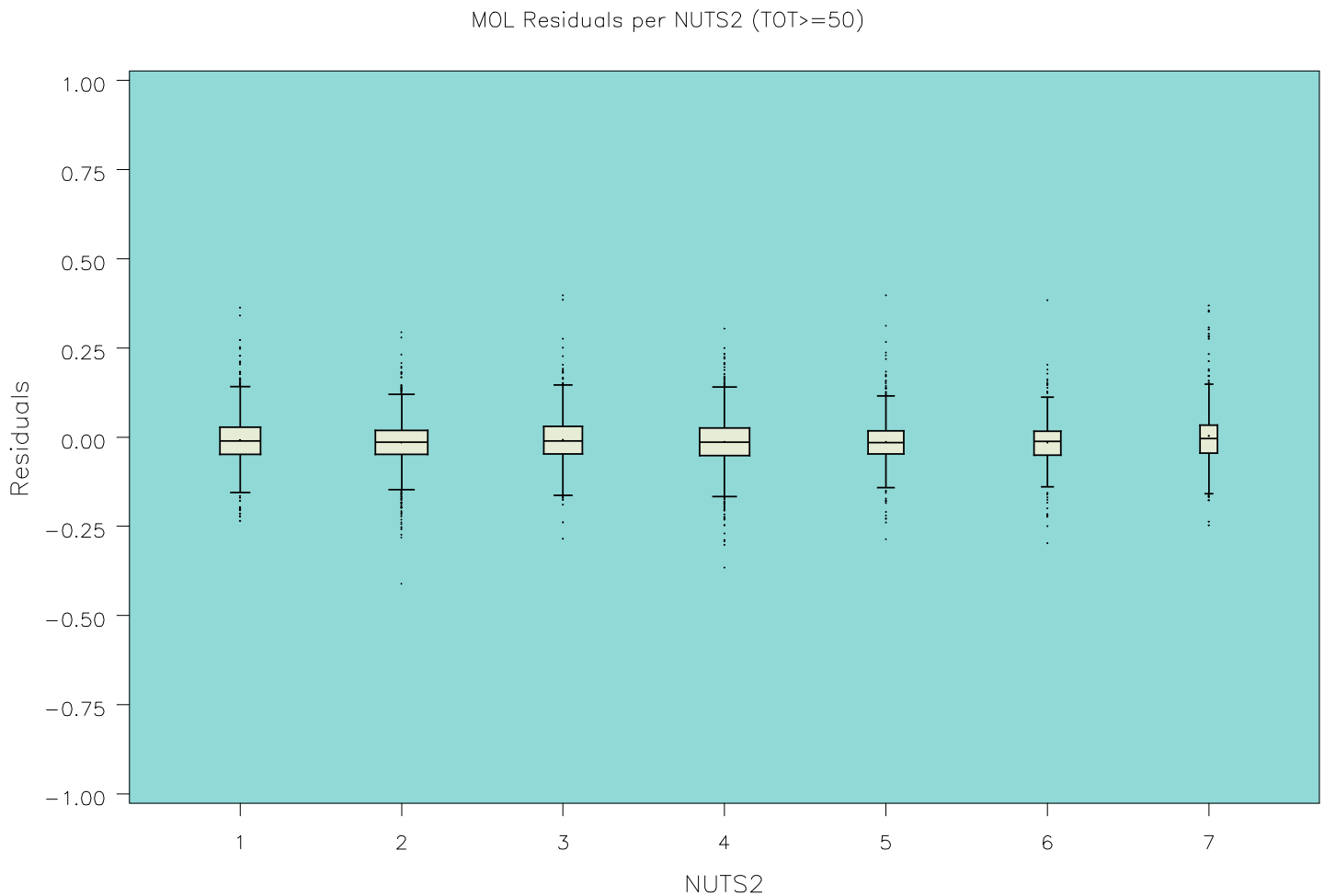

| | $R^2_{adj}$ | min (%) | max (%) |
|---|---|---|---|
| Model 1 | 0.72-0.84 | -6.0 | 7.6 |
| Model 2 | 0.80 | -4.6 | 4.8 |

min , max: relative difference (%) with benchmark at the NACE2 level.

# Model 2: log(MOL) residuals by NUTS2



# Model 2: log(MOL) residuals by NUTS2



# Discussion


- Prediction via the mean occupancy level permits to make full use of the provided information (total employment present in the register).
- A FTE predictive model seems to be a sensible approach to the problem of mass imputation of full time equivalents at the micro-level, when wage information is provided.
- Distinct models applied to 4 groupings do not bring any significant improvement over a unique model.

# Discussion II


- Investigations using truncated wages for the computation of the wage distribution characteristics show similar results as those presented here.
- If instead of a linear relationship we use a spline for MW, SdW and SkW, model 1 is slightly better fitting, but model 2 is not.
  We think that the additional complexity brought by splines is not worth the trouble.
- Instead of the log transformation for MOL, one could think of a logistic, but the observations in the SESS show that some enterprises report a FTE larger than TOT.