

Integrating households income microdata in the estimate of the Italian GDP

Alessandra Coli

Department of Statistics and Mathematics applied to Economics, University of Pisa
a.coli@ec.unipi.it

Francesca Tartamella

National Accounts Office, Istat - tartamel@istat.it

National accounts statistics are the result of the integration of several data sources. At present, in Italy, sample surveys data on households income are not included in the estimation process of national accounts aggregates. We investigate the possibility of using such data within an independent estimate of GDP, based on the income approach. The aim of this paper is to assess whether (and to what extent) sample survey microdata on household income may contribute to the estimate of GDP. To this end, surveys variables are recoded and harmonized according to national accounting concepts and definitions in order to point out discrepancies or similarities with respect to national accounts estimates. The analysis focuses particularly on compensation of income.

Keywords: reconciling micro and micro data, data integration, income approach, Gross domestic product

1. Introduction

Gross Domestic Product (GDP) can be measured according to three different methods: the production (or value added) approach, the expenditure approach and the income approach. The first method measures GDP as the value of goods and services produced by the nation, net of intermediate consumption. The expenditure approach estimates GDP in terms of the total amount of money spent for final uses of goods and services. The income approach focuses on GDP as the amount of income paid by the production units for the use of productive factors (work and capital). Due to the use of independent data sources and methods, the three approaches produce different estimates of GDP. A reconciliation process is then required in order to obtain the best estimate (that is, the published figure)¹. This reconciliation is usually carried out through the construction of Input-Output (or Supply and Use) tables. The three different approaches rely on the use of, as far as possible, independent sources of information. In the Italian national accounts, only the production and expenditure methods lead to exhaustive and independent estimates of GDP, the income approach providing adequate estimates only for some income components. This depends mainly on the insufficient quality of available data sources (Istat (2004)).

¹ The simultaneous publication of different measures is certainly useful but also misleading. On one hand it helps understanding how reliable figures are, reducing the risk of excessively trusting an inaccurate estimate. On the other hand the user has to choose among the three GDP estimates.

The paper addresses the use of sample surveys on households budgets for a better application of the income approach. As of 2004, Italy and other European countries can take advantage of a new and very rich survey, the European survey on income and living conditions (Eu-silc). The introduction of this kind of data in the GDP estimation process would allow to analyse income by household typology. Moreover it would improve reconciliation of micro and macro data on income, which is an essential piece of information for sound micro-founded macroeconomic modelling.

2. The income approach

The income approach focuses on income paid and earned by individuals and corporations in the production of goods and services, and can be represented by the following equation:

$$GDP = WS + GOS + MI + NIT \quad (1)$$

Where:

- WS: compensation of employees (wages and salaries and employers' social contributions).
- GOS: gross operating surplus, i.e. the profit generated by corporations and quasi corporations (public and private); it also includes imputed rents of owner-occupied dwellings
- MI: mixed income, i.e. the operating surplus of unincorporated enterprises owned by households. It also includes actual rents.
- NIT: taxes on production and imports net of any subsidies on production

We wonder whether it is possible (and to what extent) to estimate the terms in the right hand side of equation (1) on the basis of the households budgets surveys. Particularly, we consider the Bank of Italy survey on income and wealth (Shiw) and the European survey on income and living conditions (Eu-silc). Both surveys collect data on the different sources of household income, so it is in principle possible to estimate compensation of income (WS), mixed income (MI) and part of operating surplus (GOS). On the basis of the last five years evidence, these aggregates correspond approximately to 62% of GDP net of taxes and subsidies on production and imports.

3. How much do survey microdata fit National accounts?

Before thinking of any method for integrating surveys microdata in national accounts, it is essential to verify how much survey microdata fit national accounts estimates, because macroscopic differences would rule out the use of sample surveys microdata in national accounts. For this purpose, we estimated an almost complete Income account on the basis of Shiw and Eu-silc data sets. Surveys variables have been fully harmonised to national accounts with respect both to the observed population (i.e. Households) and to the content of each income component (Coli, Tartamella (2008)). Results are then compared with national accounts² in Tables 1 and 2.

² To compare data correctly it is necessary to estimate national accounts income components *net* of taxes and social contributions. We assume a proportional taxation on all income components. Actually, a microsimulation model should be used to correctly allocate taxation among different income categories.

Table 1 *NA and Shiw data on Households disposable income - Italy, 2004 -*

	2004*		
	NA	SHIW	SHIW/NA
Wages and salaries, net of social contributions paid by employees	318.457	267.962	84,14%
Mixed income, net of social contribution paid by self employed)	141.489	117.913	83,34%
Gross operating surplus	79.680	126.393	158,63%
Withdrawals from the income of quasi-corporations	50.239	15.415	30,68%
Social benefits	222.855	157.987	70,89%
Dividends + other income from corporations	67.268	16.994	25,26%
Interests (received-paid) not corrected for FISIM	30.448	3.621	11,89%
Other property income and transfers (net)	6.628	5.513	83,18%
Disposable income (interests not corrected for FISIM)	917.064	711.798	77,62%

*data on income are in current million euros

Table 2 *NA and Eusilc data on Households disposable income - Italy, 2004 -*

	2004*		
	NA	EUSILC	EUSILC/NA
Wages and salaries, net of social contributions paid by employees	318.457	308.192	96,8%
Self employed income	240.066	138.835	57,8%
Gross operating surplus	79.680	131.286	164,8%
Social benefits	222.855	199.378	89,5%
Other capital income	65.408	10.818	16,5%
Other transfers (received)	21.291	5.305	24,9%
Other transfers (paid)	-30.692		
Other property income and transfers (net)			
Disposable income (interests not corrected for FISIM)	917.065	793.814	86,6%

*data on income are in current million euros

Surveys underestimate (or National accounts overestimate) all the income components with the exception of Gross operating surplus. The underestimation is particularly strong for property income and transfers. Since sources have been harmonized we conclude that differences are mainly due to selection bias and under-reporting of certain types of income.

The aim of this paper is to assess the potential contribution of the surveys to the estimate of compensation of employees, mixed income and gross operating surplus. For what concerns these components, our main findings are as follows:

- Wage and salaries is the component which is best covered by both surveys, especially by Eu-silc.
- The coverage of mixed income can be evaluated only for the Shiw, since Eusilc does not allow to disentangle mixed income from other kind of property income. The Bank of Italy survey seems to ensure a good coverage (83% in 2004).
- Gross operating surplus is overestimated by both surveys. Probably households overvalue their property or identify as an imputed rent what is an actual rent³.

³ If we consider the Shiw imputed rents on secondary dwellings as actual rents, the coverage of mixed income rises up to almost 90% (since actual rents are part of mixed income), and over-coverage of operating surplus considerably decreases to 106%.

Obviously it also is possible that national accounts dramatically underestimate imputed rents.

Coming to the wages and salaries component, Table 3 compares the number of employees and the wages and salaries per capita values as they result in national accounts (NA), Shiw and Eusilc.

Table 3 *Employees and Wages and salaries by Economic activity, Italy, 2004*^{**}

Economic activity (NACE Rev. 1 classification)	National accounts		Shiw		Eusilc	
	Employees (thousand)	WS (per-capita)*	Employees (thousand)	WS (per-capita)*	Employees (thousand)	WS (per-capita)*
Agriculture	493	10.808	842	9.356	659	9.642
Industry	5.527	16.898	6.333	14.686	5.669	16.380
<i>Industry (without construction)</i>	4.367	17.592	5.054	14.874	4.422	16.934
<i>Construction</i>	1.160	14.287	1.279	13.944	1.247	14.416
Services	12.009	18.300	10.348	16.016	10.912	16.380
<i>Trade, hotel and restaurants</i>	2.542	17.458	2.360	13.369	2.362	13.333
<i>..Transport storage and comm.</i>	1.026	25.359	722	16.358	919	19.397
<i>..Financial intermediation</i>	535	30.772	678	22.404	583	23.930
<i>..Other services for business activities</i>	1.713	16.663	640	16.424	1.031	15.696
<i>..Other services (Public admin., educat., etc)</i>	6.194	16.851	5.948	16.253	6.017	17.775
Total economy	18.029	17.665	17.523	15.215	17.240	16.567

*Current thousand euros; ** Eu-silc contains several missing values for the Economic activity. This is the reason why wages and salaries calculated in this table do not correspond to the aggregate in tables 2.

We observe a general better fit of Eu-silc estimates to national accounts data. Absolute differences computed for economic sector are on average smaller with respect to Shiw, both for per capita values, number of employees and wages and salaries. “Industry (without Construction)” and “Other services” are the best covered sectors. This is not a surprise since these sectors are only marginally effected by irregular employment. The worst results appear for “Other services for business activities”, one of the economic sectors with the highest irregular jobs rate. Unexpectedly, the number of employees captured by both surveys is significantly higher for Agriculture. Coming to wages and salaries per capita values, we observe generally lower estimates in surveys but differences are not dramatic. Our analysis, though preliminary, suggests that surveys data (Eu-silc in particular) would provide valuable information for the computation of Compensation of income within an income approach for the estimate of GDP.

References

- Brandolini, A., (1999). The Distribution of Personal Income, in Post-War Italy: Source Description, Data Quality, and the Time Pattern of Income Inequality, *Temi e discussioni*, Banca Italia.
- Coli A., Tartamella F. (2008) Income and consumption expenditure by households groups in National accounts, Iariw Conference, Portoroz, Slovenia, August 2008
- Istat (2004) Metodologie di stima degli aggregati di contabilità nazionale a prezzi correnti - *Metodi e norme* n. 21, Istat.
- Reid David J. (1968) Combining Three Estimates of Gross Domestic Product, *Economica, New Series*, Vol. 35, No. 140, pp. 431-444.